

Rekursive Substrukturkondensation bei der Methode der Finiten Elemente

Dissertation

**zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
der Fakultät für Mathematik
der Ludwig-Maximilians-Universität
München**

**vorgelegt von
Gustav Löbel
aus Landsberg/Lech**

eingereicht im Juni 1985

Rekursive Substrukturkondensation bei der Methode der Finiten Elemente

Dissertation

zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
der Fakultät für Mathematik
der Ludwig-Maximilians-Universität
München

vorgelegt von

Gustav Löbel

aus Landsberg/Lech

eingereicht im Juni 1985

1. Berichterstatter: Prof. Dr. A. Sachs

2. Berichterstatter: Prof. Dr. E. Wienholtz

Tag der mündlichen Prüfung: 18. Februar 1986

INHALT

REKURSIVE SUBSTRUKTURKONDENSATION

BEI DER METHODE DER FINITEN ELEMENTE

EINLEITUNG

Problemstellung	1
Zusammenfassung	7

KAPITEL 1: SUBSTRUKTURKONDENSATION

1) Die Methode der finiten Elemente	8
2) Statische Kondensation	16
3) Kontinuierliche Substrukturkondensation	22

KAPITEL 2: REKURSIVE SUBSTRUKTURIERUNG

4) Verträglichkeit	32
5) Rekursive F.E.-Zerlegung	45
6) Rekursionsformel	54
7) Eigenschaften der Rekursion	59

KAPITEL 3: DIE FIXPUNKTGLEICHUNG

8) Matrixpolynome und latente Wurzeln	70
9) Lösungsverfahren	81
10) Bedeutung der latenten Wurzeln	88

KAPITEL 4: KONVERGENZ UND NUMERISCHE ERGEBNISSE

11) Konvergenz bei Singularitäten und unbeschr. Geb..	95
12) Numerische Ergebnisse	107

ANHANG: Literatur

Symbolverzeichnis

Problemstellung:

Viele Probleme aus der mathematischen Physik und aus ingenieurtechnischen Anwendungen führen auf Differentialgleichungen, deren Lösung in der Nähe bestimmter Punkte ein asymptotisches Verhalten der Form $O(r^s)$ zeigen. Dabei ist r der (euklidische) Abstand zu dem fraglichen Punkt und $0 < s < 1$ ein problemabhängiger Exponent.

Beispiele sind etwa der Spitzeneffekt in elektrischen Feldern, oder die Spannungskonzentrationen an scharfen Kerben in mechanisch belasteten Bauteilen.

Mathematisch gesehen entstehen solche Singularitäten an den einspringenden Ecken des Grundgebietes bei elliptischen partiellen Differentialgleichungen. (LEHMANN (1957), KONDRATJEV (1967), KELLOG (1971)). Nach KONDRATJEV läßt sich die Lösung u an solchen Eckpunkten bzw. "konischen Punkten" im $\mathbb{R}^n$ darstellen in Form einer Reihe:

$$u = \sum_{j=1}^{\infty} \sum_{k=0}^{k_j} r^{s_j} \ln^k(r) \psi_{jk}(\omega_1, \dots, \omega_{n-1})$$

wobei $\omega_1 \dots \omega_{n-1}$ und r n -dimensionale Polarkoordinaten sind. Die ψ_{jk} sind glatte Funktionen und die Zahlen s_j , k_j hängen nur noch vom Hauptteil der Gleichung und den Randbedingungen ab. Die logarithmischen Anteile stammen dabei aus der Inhomogenität der Gleichung.

So lassen sich beispielsweise in der ebenen, linearen Elastizitätstheorie die Verschiebungen an einer Rißspitze (= einspringende Ecke mit Innenwinkel 2π) in eine Reihe entwickeln, die mit

Termen der Form

$$u(r, \omega) = K r^{1/2} (2G\pi)^{-1/2} \cos(\omega/2) (1 - 2\nu + \sin^2(\omega/2)) + \dots$$

beginnt. G und ν sind Materialkonstanten. Der Koeffizient K des ersten Termes wird dabei als "Spannungsintensitätsfaktor" bezeichnet und in der Theorie der Bruchmechanik als Maß für die Gefährlichkeit eines Risses verwendet (IRWIN (1958)).

Eng verwandt mit diesem Phänomen ist das sogenannte Interfaceproblem, das vorliegt, wenn mehrere Bereiche mit unterschiedlichen (Material-) Koeffizienten aneinandergrenzen. Hier gilt eine ganz ähnliche Reihenentwicklung (KELLOG (1971), (1975)).

Für Probleme aus der ebenen Elastizitätstheorie gibt es eine große Anzahl von Arbeiten zu diesem Thema. Insbesondere das Problem eines Risses, der an einem Interface endet, wurde ausführlich untersucht. (BOGY (1968), (1971a), (1971b), ENGLAND (1965), WILLIAMS (1952), THEOCARIS (1974), BOGY, WANG (1971)). Dabei können die Exponenten s_j sogar komplex werden. Eine ausführliche Diskussion und Literaturstudie zu diesem Problemkreis findet man bei LÖBEL, ZEILINGER, DESKA (1984). Dort wird die Adaption der klassischen Bruchmechanik auf moderne Faserverbundwerkstoffe untersucht.

Die Methode der Finiten Elemente, die sich für die Lösung partieller Differentialgleichungen in der Praxis sehr stark etabliert hat, führt bei solchen Singularitäten aber meist zu völlig unbefriedigenden Ergebnissen.

Das liegt daran, daß die (üblicherweise) polynomialen Ansatzfunktionen bei gleichmäßiger Elementeinteilung die unbeschränkten Ableitungen an den Singularitäten nur ungenau modellieren können und so ein großer lokaler Fehler entsteht. Dieser lokale Fehler klingt zwar mit wachsendem Abstand rasch ab, reicht aber meist aus, um die Näherungslösung nachhaltig zu verschmutzen. ("Pollu-

tion-effect")

Dieser Effekt zerstört nicht nur die lokale Genauigkeit des Verfahrens, sondern vor allem auch dessen globales Konvergenzverhalten. Dieses hängt ja, wie man weiß (z.B. STRANG, FIX (1973), CIARLET (1978)), nicht nur von der Art der verwendeten Elemente, sondern auch von der Glattheit der wahren Lösung ab. So erreicht man bei den geschilderten Problemen bei gleichmäßiger Elemententeilung je nach Wert von s_1 völlig indiskutable Konvergenzordnungen.

Die einfachste Abhilfe ist natürlich eine lokale Verfeinerung des Elementnetzes an den kritischen Stellen. SCHATZ, WAHLBIN (1978), geben allgemeine Fehlerabschätzungen in der Maximumnorm für diese Fälle und zeigen im 2. Teil ihrer Arbeit (1979), daß man durch geeignete lokale Verfeinerung der Elemententeilung durchaus die maximal mögliche Konvergenzordnung retten kann. In der Praxis setzt diese Vorgehensweise aber ein hohes Maß an Erfahrung beim Anwender voraus, macht viel Arbeit und findet meist an der begrenzten Rechnerkapazität rasch Grenzen.

Ein anderer Weg die lokale Approximationsgüte zu erhöhen, ist die Verwendung spezieller Ansatzfunktionen. In der Literatur finden sich insbesondere wieder für die Bruchmechanik eine Vielzahl von Veröffentlichungen. Prinzipiell kann man dabei 2 Varianten dieser Idee unterscheiden:

- 1.) Wenn die Form der analytischen Lösung bekannt ist, kann man die ersten Terme der asymptotischen Entwicklung als Ansatzfunktionen verwenden. Die Schwierigkeit dabei ist, diese meist sehr komplizierten Funktionen in vernünftiger Weise in das Gesamtkonzept einzubauen (BYSKOV (1984), HEPPLER, HANSEN (1981), PAPAIONNOU, HILTON (1974), WALSH (1971)).
- 2.) Wenn wenigstens noch die Ordnung der Singularität bekannt ist, kann man mit unspezifischen Funktionen auskommen. Diese sind meist einfacher in der Handhabung. (STAAB, SUN (1981),

JOHNSON (1981)). Ein wichtiges Beispiel sind hierzu auch die "kollabierten Viertelpunktelemente" von BARSOUM (1976), (1977) bzw. HENSHELL, SHAW (1975).

Für Surveys zu F.E.-Anwendungen in der Bruchmechanik sei auf die Artikel von GALLAGHER (1979), WILSON (1973) und WATWOOD (1969) verwiesen.

Bei allen diesen Ansätzen ist die Kenntnis der entsprechenden Reihenentwicklung oder wenigstens der Ordnung s , des ersten Termes unentbehrlich. Speziell bei Interfaceproblemen kann aber allein schon die Berechnung von s eine nichttriviale Aufgabe sein (BOGY (1968), FENNER (1976), ERDOGAN, GUPTA (1971), KARP, KARAL (1962)).

Auch bei höherer Raumdimension ist die explizite Bestimmung der Singularität im allgemeinen nicht einfach (KELLOG (1971)).

Ein weiteres verwandtes Problem, über das in jüngerer Zeit recht viel publiziert worden ist, entsteht, wenn man Außenraumaufgaben betrachtet, wo die Lösungen nach außen ($r \rightarrow \infty$) ein Abklingverhalten der Form $O(r^s)$ zeigen. Hier ist $s < 0$. Meist wird dieses Verhalten explizit von der Lösung gefordert, um sachgemäße Probleme zu erhalten. (z.B. LEIS (1967)).

Anwendungsbeispiele sind etwa Felsspannungen in der Umgebung eines Tunnels, Antennenprobleme in der Atmosphäre oder stationäre Wärmeleitung um eine Pipe-line im Permafrostgebiet.

Die wahre Lösung hat auch hier i. allg. eine Entwicklung in Terme der Form $r^s \psi(\omega)$. (z.B.: GOLDSTEIN (1981a) Lemma 2.1)

Für die numerische Behandlung behilft man sich hier meist durch Abschneiden des (eigentlich unbegrenzten) Gebietes und Einführung künstlicher Randbedingungen auf dem so geschaffenen Rand (z.B. TAYLOR, HOOD (1973), GOLDSTEIN (1981a), (1981b), (1982), BAYLISS, GUNZBERGER, TURKEL (1982)). Speziell für die HELMHOLTZ-Gleichung gibt es eine Reihe von Arbeiten zu diesem Problem. (AZIZ, DORR,

KELLOG (1982), SARANEN (1981), CANTOR (1983), BABUSKA (1972)).

Eine andere Möglichkeit bieten auch hier spezielle finite Elemente - sozusagen "infinite finite Elemente" -, die sich bis Unendlich erstrecken und das (a priori bekannte !) Abklingverhalten mit geeigneten Ansatzfunktionen modellieren. (BEER (1983), ZIENKIEWICZ, EMSON, BETTES (1977), LYNN, HADID (1981), BEER, MEEK (1981), BETTES (1977), CURNIER (1983))

In der vorliegenden Arbeit wird ein Verfahren vorgeschlagen, das bei relativ geringem Aufwand an Computerkapazität und Manpower eine unendliche Verfeinerung oder Vergröberung der Elementeteilung simuliert und so die optimale Konvergenz der F.E.-Methode rettet. Insbesondere ist die Kenntnis von s nicht nötig, im Gegenteil, das vorgeschlagene Verfahren kann auch dazu benutzt werden Näherungen für s zu berechnen.

Diese unendliche Verfeinerung/Vergröberung muß aber nicht tatsächlich ausgeführt werden, sondern die Grenzgleichung, auf die sie führt wird mittels einer Eigensystemanalyse direkt simuliert. Auf diese Weise erhält man ein Macro-element, das das $O(r^s)$ -Verhalten ausgezeichnet modellieren kann.

Ganz neu ist die Grundidee allerdings nicht. Bereits 1975 hat THATCHER einen ähnlichen Ansatz verfolgt, bei dem er eine spezielle unendliche Verfeinerung für das MOTZ-Problem (MOTZ (1946)) betrachtet. Zur Auflösung des entstehenden unendlichen Gleichungssystems verwendet er jedoch eine Art Mehrschrittverfahren, das zu teilweise sehr restriktiven Einschränkungen zwingt, die genau genommen nur bei einfachsten Elementen erfüllt sind. In seiner 2. Arbeit wird dies sogar noch weiter spezialisiert, kann dann dafür dann weitgehend analytisch aufgelöst werden (THATCHER (1976)).

1982 erschien eine Arbeit von HOUDÉ HAN, in der die Methode der unendlichen Gitterverfeinerung auf 2-dimensionale Interfaceprobleme angewendet wird. Auch er betrachtet nur eine spezielle Art

der Gitterverfeinerung, verwendet aber zur Auflösung des Systems bereits eine Variante der hier vorgeschlagenen Substrukturtechnik, allerdings ohne expliziten Konvergenzbeweis.

Es zeigt sich, daß der in der vorliegenden Arbeit eingeschlagene Weg der "rekursiven Substrukturkondensation" erheblich tragfähiger ist, als diese Ansätze. So ist es damit möglich, das Verfahren zu einer allgemeinen Methode zu abstrahieren, die nicht nur für andere Differentialgleichungen und höhere Raumdimensionen sondern vor allem für (fast) beliebige Elemente anwendbar ist.

Insbesondere werden die Voraussetzungen in allgemeiner Form entwickelt, ein Konvergenzbeweis für das Verfahren angegeben und verschiedene effiziente Lösungsmethoden für das entstehende unendliche lineare Gleichungssystem vorgestellt.

Die Anwendung auf Außenraumprobleme scheint in der amerikanisch-europäischen Literatur noch neu zu sein. Nur in einer chinesischen Arbeit (SHAO XIU MING, WANG SHU JING (1983)), von der bis jetzt aber leider nur das Abstract zugänglich ist, scheinen ähnliche Ideen verfolgt worden sein.

An dieser Stelle möchte ich mich ganz besonders herzlich bei Herrn Prof. Dr. A. Sachs für die Anregung und Betreuung der vorliegenden Arbeit bedanken.

Besonderer Dank gebührt auch meinen Kollegen Dr. E. Rank und Dipl.-Ing. W. Sichert für die Durchsicht des Manuskripts.

Herrn Prof. Dr. G. Kuhn bin ich zu Dank verpflichtet für sein Interesse und die Erlaubnis die technischen Möglichkeiten seines Instituts zu benutzen.

Besonders tief stehe ich aber bei meiner Frau Kristine und meinem Sohn Benny in der Schuld, deren Geduld und Toleranz die Arbeit ermöglicht haben.

ZUSAMMENFASSUNG

In §1 wird ein kurzer Abriss der Methode der Finiten Elemente gegeben, um wichtige Begriffe bereitzustellen. Insbesondere wird der Standardtrick der statischen Kondensation ausführlich dargestellt (§2). Da er das Fundament der ganzen Arbeit ist, wird in §3 das zugrundeliegende Prinzip der Abspaltung von Teilgebieten als abstraktes Hilbertraum-Problem näher analysiert.

Die entscheidende Beobachtung dieser Arbeit ist, daß sich die Steifigkeitsmatrix eines Finite-Elemente-Ansatzes nicht bzw. nur sehr einfach ändert, wenn man das Grundgebiet bestimmten affinen Transformationen unterwirft. In §4 wird dieser Sachverhalt genauer dargestellt und die Verträglichkeit einer Bilinearform mit einer affinen Transformation definiert.

§5 diskutiert dann die Idee N transformierte Kopien einer Grundstruktur miteinander zu koppeln. Durch geeignete Wahl der Grundstruktur und der affinen Transformation kann man auf diese Weise insbesondere Gitterverfeinerungen und Vergrößerungen erzielen.

Durch sukzessive statische Kondensation der Substrukturen erhält man eine Rekursionsformel, die in §6 hergeleitet wird und deren Konvergenz für $N \rightarrow \infty$ gegen einen Fixpunkt in §7 gezeigt wird. Die Fixpunktgleichung ist ein quadratisches Matrixpolynom, dessen Eigenschaften in §8 näher untersucht werden. §9 diskutiert dann verschiedene Möglichkeiten diese Gleichung direkt zu lösen und §10 zeigt die anschauliche Bedeutung der Fixpunktmatrix auf.

Damit wird es möglich unendliche Gitterverfeinerungen/ -vergrößerungen zu simulieren, ohne sie wirklich auszuführen. Für praktische Anwendungen ist dies insbesondere deswegen interessant, weil man damit die optimale Konvergenzordnung der Finite-Elemente-Methode auch bei Singularitäten und unbeschränkten Gebieten retten kann. in §11 werden die diesbezüglichen Konvergenzbeweise formuliert und §12 gibt schließlich numerische Ergebnisse.

KAPITEL 1: FINITE ELEMENTE UND SUBSTRUKTURKONDENSATION

1) DIE METHODE DER FINITEN ELEMENTE

Um die Notation zu fixieren und einige wichtige Begriffe zusammenzustellen, soll hier kurz die Methode der finiten Elemente skizziert werden. Für ausführlichere Beschreibungen sei auf die Literatur verwiesen (CIARLET (1978), ZIENKIEWICZ (1977), STRANG, FIX (1973)). Wir stützen uns hier vor allem auf die Darstellung von CIARLET.

Das Problem möge in Form einer Variationsaufgabe folgender Gestalt vorliegen:

$\Omega \subset \mathbb{R}^n$ Grundgebiet, V Hilbertraum aus reellwertigen Funktionen auf Ω

$a: V \times V \longrightarrow \mathbb{R}$ stetige, V -elliptische Bilinearform)*

$f: V \longrightarrow \mathbb{R}$ stetiges lineares Funktional

gesucht: $u \in V$ mit $\forall v \in V : a(u, v) = f(v)$ (1.0)

Nach dem Lemma von LAX-MILGRAM ist die Aufgabe in dieser Form sachgemäß gestellt - also insbesondere auch eindeutig lösbar.

*) d.h. Es gibt Konstanten C und M so, daß

$\forall v \in V : |a(v, v)| \geq C \|v\|^2$ (V -elliptisch)
und $\forall u, v \in V : |a(u, v)| \leq M \|u\| \|v\|$ (stetig)

Beim klassischen GALERKIN-Ansatz wird die Lösung von (1.0) approximiert, indem man das Problem auf einen endlich dimensionalen Teilraum $V_h \hookrightarrow V$ beschränkt:

gesucht: $u_h \in V_h$ mit $\forall v \in V_h : a(u_h, v) = f(v)$ (1.1)

Mit einer Basis $\{\psi_1, \dots, \psi_m\}$ von V_h läßt sich die Näherungslösung u_h dann darstellen als Linearkombination

$$u_h(x) = \sum_{i=1}^m u_i \psi_i(x) \quad (1.2)$$

wobei sich der Vektor $U = (u_1, u_2, \dots, u_m)$ der unbekannten Koeffizienten aus folgendem linearen Gleichungssystem berechnen läßt:

$$K U = F \quad (1.3)$$

$$\text{mit: } K = (a(\psi_i, \psi_j))_{i,j=1 \dots m} \quad (1.3.a)$$

$$F = (f(\psi_i))_{i=1 \dots m} \quad (1.3.b)$$

In Anlehnung an die anschauliche physikalische Bedeutung bei elastostatischen Problemen wird K meist als "Steifigkeitsmatrix", F als "Lastvektor" und U als "Verschiebungsvektor" oder Vektor der Knotenverschiebungen bezeichnet.

Das wesentliche an der Methode der finiten Elemente besteht nun in der speziellen Wahl des Teilraumes V_h . Dieser wird aus einzelnen finiten Elementen zusammengebaut. So erreicht man, daß die Ansatzfunktionen lokalen Träger haben und damit die Steifigkeitsmatrix nur in einem dünnen Band um die Hauptdiagonale besetzt ist. Der Index h bezieht sich dabei auf eine charakteristische Länge der Elementzerlegung.

Nach CIARLET (1978) ist dabei ein allgemeines finites Element wie folgt definiert:

(1.1) DEFINITION:

Ein finites Element ist ein Tripel (E, P, Σ) mit:

- (i) E ist eine abgeschlossene Teilmenge des $\mathbb{R}^n$ mit nicht-leerem Inneren und LIPSCHITZ-stetigem Rand
- (ii) P ist ein Raum von reellwertigen Funktionen definiert auf E
- (iii) Σ ist ein Satz von endlich vielen, linear unabhängigen Linearformen ϕ_i $i=1, \dots, k$ definiert auf P , die P -unisolvant sind.)*

Meist wird etwas salopp auch die Teilmenge E allein schon als finites Element bezeichnet. In aller Regel nimmt man hierfür einfache geometrische Figuren, wie Dreiecke oder Quader.

P ist der Raum der Ansatzfunktionen. Normalerweise verwendet man hier Polynome, aber auch andere Ansätze sind möglich (z.B. bei speziellen Rißspitzenelementen).

Σ heißt Menge der Freiheitsgrade (degrees of freedom, kurz: dof's) und enthält meist Punktfunktionale vom Typ

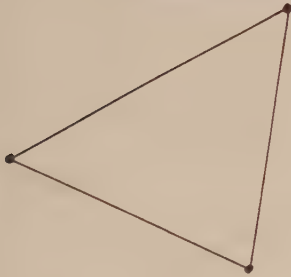
$$\text{"Funktionsauswertung"} \quad \phi_i(p) = p(x_i) \quad (1.4.a)$$

$$\text{oder "Richtungsableitung"} \quad \phi_i(p) = \frac{\partial p}{\partial \xi}(x_i) \quad (1.4.b)$$

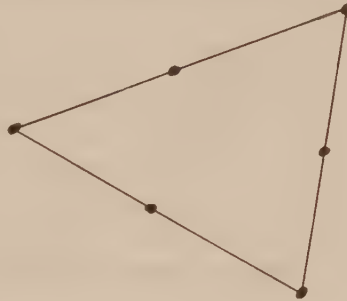
Auch Funktionale mit höheren Ableitungen oder Integralen sind gebräuchlich, wir wollen diese Details hier aber nicht zu weit verfolgen. Die Punkte $x=x_i$ $i=1, \dots, k$ werden dabei als Knoten bezeichnet.

*) d.h. zu allen k -tupeln α_i $i=1, \dots, k$ gibt es genau ein $p \in P$ mit $\phi_i(p) = \alpha_i$ für alle $i=1, \dots, k$.

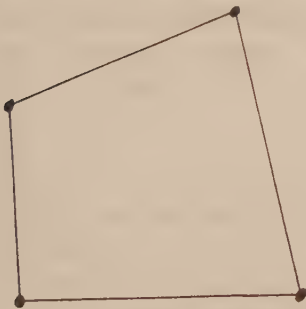
Das einfachste und wohl bekannteste 2 dimensionale finite Element besteht aus einem Dreieck als Grundgebiet E , über dem lineare Polynome $p(x,y) = a + bx + cy$ betrachtet werden. Als Freiheitsgrade hat es die Funktionswerte an den Ecken. In der Literatur ist es unter den verschiedensten Namen bekannt: TRIM3, PSI3, Lineares Dreieck, COURANT-Element, Constant Strain Triangle, usw.



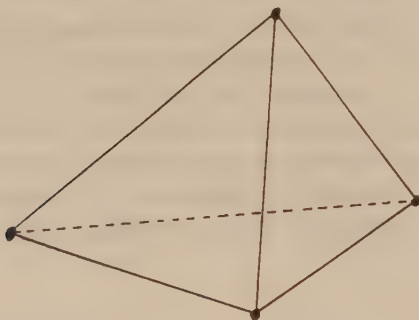
Das nächstkompliziertere 2d-Element ist ein Dreieck mit quadratischem Polynomansatz $q(x,y) = a + bx + cy + dx^2 + exy + fy^2$. Seine 6 Freiheitsgrade - Punktfunktionale vom Typ (1.4.a) - werten an den Ecken und den Seitenmitten aus. Es wird z.B. als TRIM6, PSI6 oder quadratisches Dreieck bezeichnet.



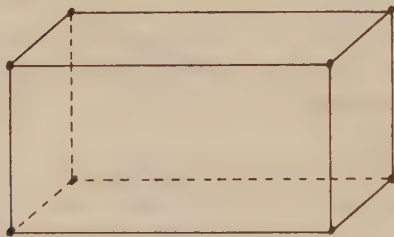
Ein weiteres in der Praxis sehr häufig verwendetes Element ist das "isoparametrische Viereck". Seine 4 Freiheitsgrade sind die Funktionswerte an den Ecken. Sein Ansatz geht aus einem unvollständigen Polynom 2. Grades auf dem Einheitsquadrat hervor, das dann durch isoparametrische Transformation entsprechend verzerrt wird. Ganz analog wie bei den Dreiecken gibt es dazu auch eine Variante mit Seitenmittenknoten und also 8 Freiheitsgraden.



Von den 3-dimensionalen Elementen sollen hier nur der Vollständigkeit halber noch kurz die Familie der Tetraederelemente



und die Familie der isoparametrischen Quaderelemente erwähnt werden.



Durch geeignetes Zusammensetzen mehrerer finiter Elemente (E_i, P_i, Σ_i) $i=1, \dots, e$ wird dann der Approximationsraum V_h erzeugt. Dazu wird das Grundgebiet Ω durch eine Vereinigung von paarweise disjunkten Elementen ausgeschöpft:

$$\Omega = \bigcup_{i=1}^e E_i \cup R \quad (1.5)$$

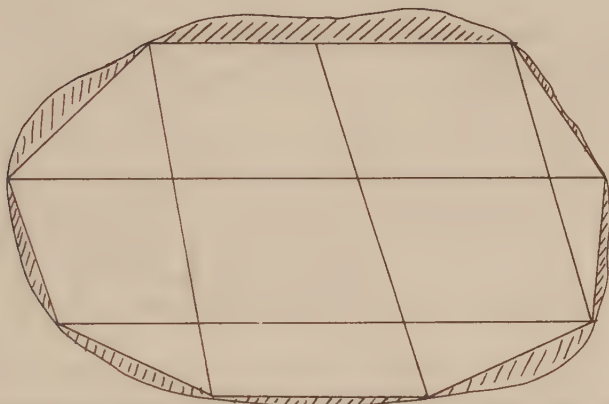
Das nicht approximierbare Restgebiet R wird vernachlässigt.

Die Räume der Ansatzfunktionen und die Freiheitsgrade werden ebenfalls zusammengefügt:

$$P_{ges} = \{ f : \Omega \longrightarrow \mathbb{R} \mid f|_{E_i} \in P_i \}$$

$$\Sigma_{ges} = \bigcup_{i=1}^e \Sigma_i \quad (1.6)$$

Skizze:

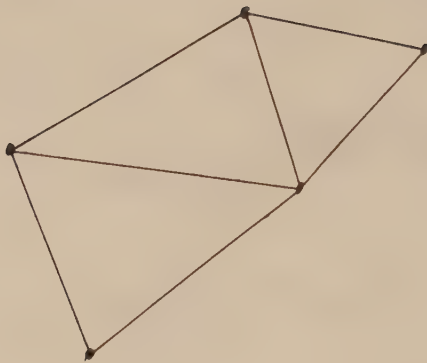


Damit der so gewonnene Raum $V_h = P_{ges}$ auch tatsächlich ein sinnvoller Teilraum von V wird, muß noch folgende Konnektivitätsbedingung erfüllt sein:

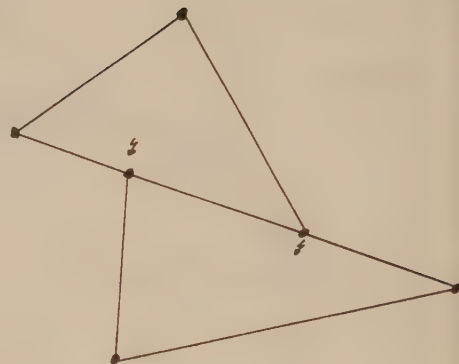
- (i) Σ_{ges} ist P_{ges} -unisolvant
- (ii) $P_{ges} \subset V$
- (1.7)

In den meisten Fällen kann man diese formale Bedingung durch die einfachere, bekannte Forderung ersetzen, daß benachbarte Elemente nur ganze Kanten oder Ecken gemeinsam haben dürfen und auch die Freiheitsgrade längs der gemeinsamen Kante aneinanderpassen müssen.

Skizze:



erlaubt



nicht erlaubt

Man sieht sofort, daß auch jede Zusammenfügung von finiten Elementen wieder die Definition eines finiten Elementes erfüllt. Wir wollen dies zur Unterscheidung ein Macro-Element nennen. Ein Macro-Element, das das gesamte Grundgebiet Ω ausfüllt (oder wenigstens einigermaßen approximiert), wollen wir eine Finite-Elemente-Zerlegung (F.E.-Zerlegung) von Ω nennen.

Mit einer F.E.-Zerlegung definiert man nun in kanonischer Weise eine Basis $\{\psi_1, \dots, \psi_m\}$ von V_h durch die Bedingung

$$\phi_i(\psi_j) = \delta_{ij} \quad \text{für alle Funktionale } \phi_i \in \Sigma_{\text{ges}}$$

Durch diese Basis wird über Gl. (1.3) für eine gegebene F.E.-Zerlegung in eindeutiger Weise eine Steifigkeitsmatrix definiert.

Ein für die weiteren Ausführungen eminent wichtiger Begriff ist die Affin-Äquivalenz von finiten Elementen. In Anlehnung an

CIARLET (1978) wollen wir diese folgendermaßen definieren:

(1.2) DEFINITION:

Zwei finite Elemente (E_1, P_1, Σ_1) und (E_2, P_2, Σ_2) mit Freiheitsgraden vom Typ wie in (1.4) an Knoten $x_i^{(1)}$ bzw. $x_i^{(2)}$ und eventuell Richtungsableitungen nach $\xi_k^{(1)}$ bzw. $\xi_k^{(2)}$ heißen affin äquivalent : $\langle == \rangle$

$\exists \hat{T} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ affine, bijektive Abbildung mit:

$$(i) \quad E_2 = \hat{T}(E_1)$$

$$(ii) \quad P_2 = \{ p \circ \hat{T}^{-1} \mid p \in P_1 \}$$

$$(iii) \quad x_1^{(2)} = \hat{T}(x_1^{(1)}) \quad \text{und} \quad \xi_k^{(2)} = \hat{T} \xi_k^{(1)}$$

In dieser Weise läßt sich auch definieren, wann zwei komplette F.E.-Zerlegungen zu verschiedenen Grundgebieten affin äquivalent heißen, indem man beide als Macro-Elemente betrachtet.

Wie der Name schon andeutet, definiert die Affinäquivalenz eine Äquivalenzrelation auf den finiten Elementen. Dies wird meist ausgenützt, indem man die Formfunktionen nur einmal auf einem einfachen Standardelement (Einheitsquadrat, Einheitsdreieck...) als Repräsentant der Äquivalenzklasse herleitet. Daraus werden dann durch entsprechende Transformation leicht Elemente in beliebiger Lage erzeugt.

Der Begriff der Affinäquivalenz eines Elementes wird meist im Sinne von "affin äquivalent zu einem einfachen Standardelement" gebraucht. Alle oben zitierten Elemente haben die Eigenschaft der Affinäquivalenz in diesem Sinne. Untereinander sind sie aber natürlich nicht affin äquivalent.

2) STATISCHE KONDENSATION

Im Weiteren wird der Begriff der statischen Kondensation von Substrukturen eine zentrale Rolle spielen. Dabei handelt es sich um einen Standardtrick aus der F.E.-Praxis, durch den eine beträchtliche Reduktion der Unbekannten erreicht werden kann. Die Behandlung großer Probleme wird dadurch oft überhaupt erst ermöglicht.

Große, komplex aufgebaute Strukturen, wie ganze Flugzeuge oder Automobile werden dazu in Teilbezirke (sog. "Substrukturen") zergliedert, welche selbst jeweils in finite Elemente eingeteilt werden. Diese Substrukturen kann man als Macro-Elemente auffassen, die natürlich untereinander die Konnektivitätsbedingung (1.7) erfüllen müssen.

In jeder dieser Substrukturen lassen sich dann alle inneren Knoten, die nicht in direkter Wechselwirkung mit einer Nachbarstruktur stehen, a priori eliminieren ("kondensieren"). Anschließend müssen die einzelnen Teile nur noch längs ihrer Nahtstellen gekoppelt werden.

Formal handelt es sich bei diesem Trick im wesentlichen um eine Partitionierung und blockweise Auflösung des linearen Gleichungssystems (1.3). Am Beispiel von 2 Teilstrukturen soll die Vorgehensweise kurz erläutert werden

Der Vektor U der unbekannten Knotenwerte wird dazu (eventuell nach geeigneter Umordnung) in $U = (U_1, U_2)$ partitioniert, wobei U_1 alle Knotenwerte zu inneren Knoten der ersten Substruktur und U_2 alle übrigen Werte (incl. Nahtknoten) enthält:

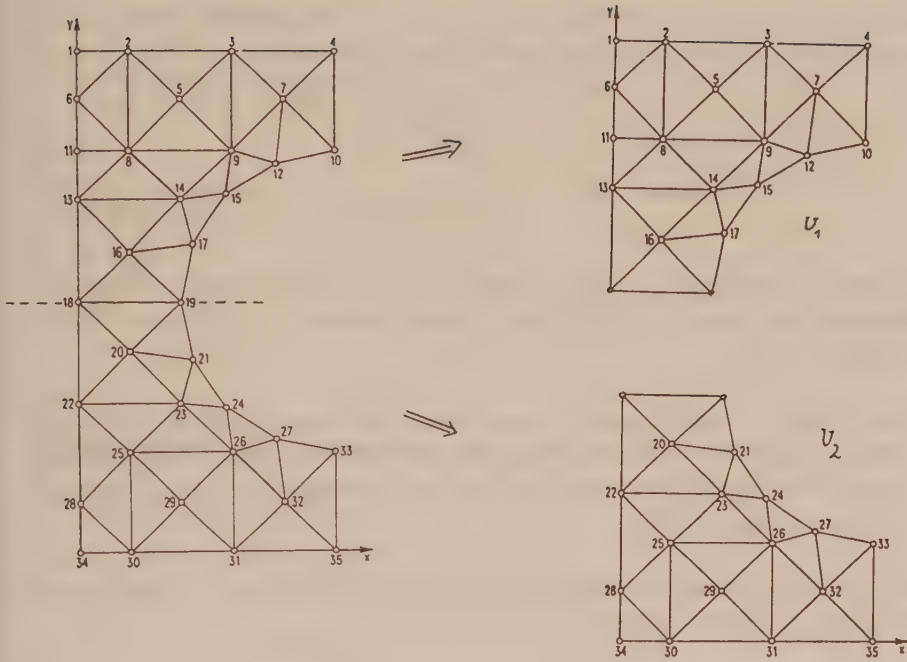


Bild 1: Substrukturzerlegung

Entsprechend werden auch die Steifigkeitsmatrix K und der Lastvektor F gegliedert:

$$\begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} \begin{matrix} \} m_1 \text{ Zeilen} \\ \} m_2 \text{ Zeilen} \end{matrix} \quad (2.1)$$

Löst man nun die 2. Blockzeile nach U_2 auf - das nachfolgende Lemma zeigt, daß das normalerweise auch geht - ,so folgt:

$$U_2 = k_{22}^{-1}(F_2 - k_{21}U_1) \quad (2.2)$$

das kann man nun wiederum in die 1. Blockzeile von (2.1) einsetzen und erhält so das kondensierte System:

$$\underbrace{(k_{11} - k_{12}k_{22}^{-1}k_{21})}_{K^*} U_1 = \underbrace{F_1 - k_{22}^{-1}F_2}_{F^*} \quad (2.3)$$

Dieses kondensierte Gleichungssystem enthält nun als Unbekannte nur noch den Vektor U_1 . U_2 kann anschließend aus Gleichung (2.2) gewonnen werden.

Man beachte, daß dieser Prozess keinerlei zusätzliche Approximation einführt. Die komplette Information der Steifigkeitsmatrix bleibt voll erhalten - es ergibt sich exakt die gleiche Lösung wie ohne Kondensation.

Daß dieser Vorgang praktisch auch immer funktioniert zeigt das folgende einfache Lemma:

(2.1) LEMMA:

Die zugrunde gelegte Bilinearform $a(.,.) : V \times V \longrightarrow \mathbb{R}$
 sei V -elliptisch.

Dann gilt:

- i) Die Steifigkeitsmatrix K in (1.3) ist positiv definit
- ii) Bei der Partitionierung gemäß (2.1) sind k_{11} und k_{22} positiv definit (also insbesondere nichtsingulär)
- iii) die kondensierte Matrix K^* gemäß (2.3) ist positiv definit.

BEWEIS:

zu i)

Sei $x = (x_1, \dots, x_m) \in \mathbb{R}^m$, $x \neq 0$ gegeben.

setze $w := \sum x_i \psi_i$ mit $\{\psi_i\}$ Basis von V_h
 $\Rightarrow x^T K x = a(w, w) > \text{Const. } \|w\|^2 > 0$

zu ii)

Sei $y \in \mathbb{R}^m, y \neq 0$, bel. vorgegeben

$$\Rightarrow y^T k_{22} y = \begin{pmatrix} 0 \\ y^T \end{pmatrix}^T K \begin{pmatrix} 0 \\ y \end{pmatrix} > 0$$

* ↑ wegen 1)

analog für k_{21}

zu iii)

Sei $y \in \mathbb{R}^m$ bel. vorgegeben, $y \neq 0$

$$\Rightarrow y^T K^* y = y^T \begin{pmatrix} k_{11} & -k_{12} & k_{12}^{-1} k_{22} & k_{21} \end{pmatrix} y =$$

$$= \begin{pmatrix} y^T \\ -(k_{22}^{-1} k_{21} y)^T \end{pmatrix}^T K \begin{pmatrix} y \\ -(k_{22}^{-1} k_{21} y) \end{pmatrix} > 0$$

q.e.d.

In dieser kompakten Schreibweise mutet die statische Kondensation beinahe so an, als könnte sie numerischen Aufwand sparen. Das täuscht jedoch!

Der Rechenaufwand bleibt praktisch der Gleiche. Im Gegenteil, durch einen gewissen zusätzlichen Verwaltungsaufwand wird in der konkreten Anwendung der Aufwand eher etwas größer.

Natürlich wird man in der Praxis die explizite Invertierung der Teilmatrix k_{22} zu vermeiden suchen. Wesentlich effektiver ist es, wenn man durch sukzessive Elimination der entsprechenden Variablen mittels GAUSS'scher Eliminationsschritte das kondensierte Gleichungssystem direkt erzeugt. Für nähere Einzelheiten zu einer derartigen Vorgehensweise sei z.B. auf das Buch von SCHWARZ (1980) verwiesen.

Aus dieser Perspektive ist das ganze Verfahren dann schon beinahe trivial - nichts anderes als eine bestimmte Reihenfolge der Schritte im GAUSS'schen Algorithmus zur Lösung linearer Gleichungssysteme.

Der Gewinn besteht lediglich darin, daß im Kernspeicher des Rechners nie die gesamte Steifigkeitsmatrix benötigt wird und bei einer Veränderung einzelner Substrukturen nicht jedesmal die komplette Matrix neu berechnet und invertiert werden muß.

Eine für die Praxis weniger relevante Variante dieser Technik, die wir hier entwickeln wollen, ist die Substrukturkondensation mit Überlappung. Bei der bereits beschriebenen Version werden die Unbekannten streng in 2 (bzw. mehr) Gruppen eingeteilt. Prinzipiell ist es aber möglich Überlappung dieser Gruppen zuzulassen.

Schematisch läßt sich das an folgendem 3-fach partitionierten System zeigen:

$$\begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \begin{bmatrix} \begin{pmatrix} u_1^o \\ u_1^u \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ u_2^o \\ u_2^u \end{pmatrix} \end{bmatrix} = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix} \quad (2.4)$$

Im Bereich der Überlappung - also bei u_1^u, u_2^o ist die Aufteilung natürlich nicht eindeutig festgelegt. Nur die Summe ist fest!

Man kann hier nun ganz analog die beiden unteren Block-Zeilen für vorgegebenes $(u_1^o, u_1^u, 0)$ nach $(0, u_2^o, u_2^u)$ auflösen und findet:

$$\begin{pmatrix} u_2^o \\ u_2^u \end{pmatrix} = \begin{pmatrix} k_{22} & k_{23} \\ k_{32} & k_{33} \end{pmatrix}^{-1} \left[\begin{pmatrix} F_2 \\ F_3 \end{pmatrix} - \begin{pmatrix} k_{21} & k_{22} \\ k_{31} & k_{32} \end{pmatrix} \begin{pmatrix} u_1^o \\ u_1^u \end{pmatrix} \right] \quad (2.5)$$

dies ergibt eingesetzt in die beiden oberen Block-Zeilen:

$$\left[\begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} - \begin{pmatrix} k_{12} & k_{13} \\ k_{22} & k_{33} \end{pmatrix} \begin{pmatrix} k_{22} & k_{23} \\ k_{32} & k_{33} \end{pmatrix}^{-1} \begin{pmatrix} k_{21} & k_{22} \\ k_{31} & k_{32} \end{pmatrix} \right] \begin{pmatrix} u_1^o \\ u_1^u \end{pmatrix} = \quad (2.6)$$

$=: K^*$

$$= \left[\begin{pmatrix} F_1 \\ F_2 \end{pmatrix} - \begin{pmatrix} k_{12} & k_{13} \\ k_{22} & k_{23} \end{pmatrix} \begin{pmatrix} k_{22} & k_{23} \\ k_{32} & k_{33} \end{pmatrix}^{-1} \begin{pmatrix} F_2 \\ F_3 \end{pmatrix} \right]$$

$=: F^*$

Man beachte allerdings, daß bei dieser Art der Aufteilung nach Lemma (2.1) zwar noch die Invertierbarkeit von $\begin{smallmatrix} k_{22} & k_{23} \\ k_{32} & k_{33} \end{smallmatrix}$ aber nicht mehr die von K^* gesichert ist.

In der Tat ist K^* singulär und das System (2.6) hat beliebig viele Lösungen. Für jede dieser Lösungen (u_1^o, u_1^u) liefert aber Gleichung (2.5) ein (u_2^o, u_2^u) derart, daß die Summe wie in Gl. (2.4) den (natürlich eindeutig bestimmten) Lösungsvektor des Gesamtsystems (2.4) ergibt. Der Rangdefekt von K^* entspricht dabei genau der Dimension der Überlappung.

Einen expliziten Beweis dieser Aussagen wollen wir hier nicht geben. Er folgt direkt aus Satz (3.1) des nächsten Abschnittes durch geeignete Spezialisierung.

3) KONTINUIERLICHE SUBSTRUKTURKONDENSATION

Viele Aussagen und Rechenansätze aus diskreten Problemstellungen finden ihre Entsprechung im Kontinuierlichen. So scheint es naheliegend auch nach dem kontinuierlichen Analogon der statischen Kondensation bei finiten Elementen zu suchen.

Dieses Analogon eröffnet die Möglichkeit Teilgebiete einer Aufgabe (z.B. unliebsame Außenbereiche, kritische Singularitäten) durch Kondensation zu umgehen. Wesentlich daran ist die Einsicht, daß der Vorgang der Matrixpartitionierung einer additiven Zerlegung des Grundraumes entspricht.

Vorgelegt sei also das Ausgangs-Problem (1.0). Der Funktionenraum V bestehe aus 2 Teilen V_1 und V_2 und gesucht ist ein Ersatzproblem, das vollständig in V_1 formuliert werden kann und dessen Lösung dort der Gesamtlösung entspricht.

Dies leistet folgender Satz:

(3.1) SATZ

Sei $V = V_1 + V_2$; V separabler Hilbertraum, V_1, V_2 Unterräume

Dann gilt:

Es gibt	$L: V_1 \longrightarrow V_2$	linearer beschränkter Operator
	$a: V_1 \times V_1 \longrightarrow \mathbb{R}$	stetige Bilinearform
	$f^*: V_1 \longrightarrow \mathbb{R}$	beschränktes lineares Funktional
	$u_f \in V_2$	eindeutig bestimmte Funktion

so, daß das Variationsproblem

gesucht $u_1 \in V_1 : \forall v \in V_1 : a^*(u_1, v) = f^*(v)$

bis auf Elemente aus $V_1 \cap V_2$ stets eindeutig lösbar ist und die Lösung $u \in V$ von (1.0) gegeben ist durch

$$u = u_1 - Lu_1 + u_f$$

BEWEIS:

mit $u = u_1 + u_2$, $u_i \in V_i$ $i=1,2$ ist (1.0) äquivalent zu dem System:

$$\textcircled{1} \quad \forall v \in V_1: a(u_1, v) + a(u_2, v) = f(v)$$

$$\textcircled{2} \quad \forall v \in V_2: a(u_1, v) + a(u_2, v) = f(v)$$

1) Definition von L :

für $u_1 \in V_1$ fest vorgegeben ist

$$g_{u_1}(\cdot) := a(u_1, \cdot) : V_2 \longrightarrow \mathbb{R}$$

eine stetige Linearform auf V_2 , denn

a) linear ist trivial

$$\begin{aligned} \text{b) } |a(u_1, v)| &\leq M \underbrace{\|u_1\|}_{=: \|g_{u_1}\|} \|v\| \quad (\text{Stetigkeit von a}) \\ &=: \|g_{u_1}\| \end{aligned}$$

ferner ist a natürlich V_2 -elliptisch, da sie sogar V -elliptisch ist

ergo: die Variationsaufgabe

$$(*) \text{ gesucht } u \in V_2 : \forall v \in V_2 : a(u, v) = g_{u_1}(v)$$

ist nach LAX-MILGRAM stets eindeutig lösbar für jedes beliebige $u_1 \in V_1$.

Dies definiert einen Operator $L : V_1 \longrightarrow V_2$
 $u_1 \longmapsto \text{Lösung von } (*)$

Es gilt:

L ist linear (trivial)

L ist beschränkt, denn nach LAX-MILGRAM gilt ja

$$\|Lu_1\| \leq 1/C \|g_{u_1}\| \leq M/C \|u_1\|$$

2) Definition von u_f :

Die Aufgabe $\forall v \in V_2 : a(u, v) = f(v)$

ist nach LAX-MILGRAM stets eindeutig lösbar. Die Lösung sei $u_f \in V_2$

für beliebig gegebenes $u_1 \in V_1$ ist somit $u_2 := u_f - Lu_1$ Lösung von ②

3) Definition von a^* und f^* :

setzt man u_2 in ① ein, so folgt:

$$\text{gesucht } u_1 \in V_1 : \forall v \in V_1 : a(u_1, v) + a(u_f - Lu_1, v) = f(v)$$

also:

$$\begin{aligned} \forall v \in V_1 : a(u_1 - Lu_1, v) &= f(v) - a(u_f, v) \\ &\underbrace{\hspace{1cm}}_{=: a^*(u_1, v)} \quad \underbrace{\hspace{1cm}}_{=: f^*(v)} \end{aligned}$$

4) Zeige a^* ist stetig:

$$\begin{aligned} |a^*(u, v)| &= |a(u, v) - a(Lu, v)| \leq \\ &\leq |a(u, v)| + |a(Lu, v)| \leq \\ &\leq M \|u\| \|v\| + M \|L\| \|u\| \|v\| = \\ &= M (1 + \|L\|) \|u\| \|v\|. \end{aligned}$$

5) zeige: a^* ist $V_1 / (V_1 \cap V_2)$ -elliptisch :

Sei $H := V_1 \cap V_2$

$$H_1 := \{v \in V_1 \mid (v, h) = 0 \quad \forall h \in H\}$$

$$H_2 := \{v \in V_2 \mid (v, h) = 0 \quad \forall h \in H\}$$

Bemerkung: im separablen Hilbertraum gilt $H_1 \perp H_2$ und auch $V = H_1 + H + H_2$ (Beweis über ein vollständiges Orthonormalsystem)

Es gilt $L_H = \text{id}$

denn die definierende Gleichung von L hat offensichtlich diese Lösung und die ist eindeutig bestimmt.

Nach Definition von L gilt:

$$\forall w \in V_1 : \forall v \in V_2 : a(Lw, v) = a(w, v)$$

es ist insbesondere $Lw \in V_2$ also

$$\forall w \in V_1 : a(Lw, Lw) = a(w, Lw)$$

$$\Rightarrow a(w - Lw, Lw) = 0 \quad \text{f. alle } w \in V_1$$

damit folgt:

$$\forall u \in V_1 : a^*(u, u) = a(u - Lu, u) = a(u - Lu, u - Lu) \geq c \|u - Lu\|^2$$

Zerlege nun $a^*(u_1, v) = f^*(v)$ in $H \oplus H_1$, sei $u_1 = \tilde{u} + h$ mit $\tilde{u} \in H_1$ und $h \in H$

zeige : $\tilde{u}$ ist eindeutig, h beliebig

$$(a) \quad \forall v \in H_1 : a(\tilde{u} - L\tilde{u}, v) = f^*(v)$$

$$(b) \quad \forall h \in H : a(\tilde{u} - L\tilde{u}, h) = f^*(h)$$

$$\text{Es ist aber } Lh=h \text{ also } L(u_1) = L\tilde{u}+h \implies u_1 - Lu_1 = \tilde{u} - L\tilde{u}$$

ergo $a(\tilde{u} - L\tilde{u}, h) = 0$ stets und $f^*(h) = 0$ stets nach Def. von f^* und so ist (b) also stets trivialerweise erfüllt, d.h. h ist beliebig.

6) Zeige a^* ist H_1 -elliptisch:

Sei $u \in H_1$:

$$\begin{aligned} |a^*(u, u)| &\geq C \|u - Lu\|^2 = C(u - Lu, u - Lu) = C(\underbrace{(u, u)}_{=0} - 2\underbrace{(u, Lu)}_{=0} + \underbrace{(Lu, Lu)}_{\geq 0}) \\ &\geq C \|u\|^2 \end{aligned}$$

*dabei ist $(u, Lu) = 0$ weil $u \in H_1$ und $Lu \in V_2$ und $H_1 \perp V_2$

Wiederum nach LAX-MILGRAM ist somit (a) eindeutig lösbar, d.h. $\tilde{u}$ eindeutig bestimmt

q.e.d.

(3.2) Bemerkungen:

(i) Falls $f(v) = 0$ für alle $v \in V_2$, so ist $u_f = 0$ und also auch $f^*(v) = f(v)$ für alle $v \in V$. u_1 und damit u sind also formal durch eine Aufgabe bestimmt, die vollständig in V_1 gestellt ist.

(ii) Wenn V sich in eine direkte Summe zerlegen läßt $V = V_1 \oplus V_2$,

so ist u sogar eindeutig bestimmt. Im diskreten Fall läßt sich das immer erreichen. In der Tat entspricht ja die Partitionierung des Vektors der unbekannten Knotenvariablen bei der diskreten Substrukturkondensierung genau der Zerlegung

$$V_1 = \text{span}\{\psi_1, \dots, \psi_{m_1}\} \quad V_2 = \text{span}\{\psi_{m_1}, \dots, \psi_{m_1+m_2}\}$$

für die Basis $\{\psi_1, \dots, \psi_m\}$ des Gesamtraumes. ($m_1+m_2=m$)

Wir haben bislang nichts über die Symmetrie von a vorausgesetzt. Die Bilinearform a^* wirkt auf den ersten Blick aufgrund ihrer Definition sehr unsymmetrisch. Daß dieser Eindruck täuscht zeigt das folgende Lemma.

(3.3) LEMMA:

Wenn a symmetrisch ist, so auch a^*

Beweis:

1.) Es gilt per Def. von L :

$$\forall v_2 \in V_2 : a(Lu, v_2) = a(u, v_2)$$

also wegen $Lv \in V_2$ für beliebiges $v \in V_1$

$$a(Lu, Lv) = a(u, Lv) = a(Lv, u)$$

=

$$a(Lv, Lu) = a(v, Lu) = a(Lu, v)$$

2.) Damit folgt sofort

$$\begin{aligned} a^*(u, v) &= a(u - Lu, v) = a(u, v) - a(Lu, v) = \\ &= a(v, u) - a(Lv, u) = a^*(v, u) \end{aligned}$$

q.e.d.

Wir wollen nun die Aussage des Satzes an einem konkreten einfachen Beispiel demonstrieren.

(3.4) BEISPIEL:

Als Beispiel für Satz (3.1) wollen wir folgende einfache Randwertaufgabe betrachten:

gesucht sei $u \in C^1(-1, +1)$ mit

$$-u'' = 2$$

$$u(-1) = 0$$

$$u(+1) = 0$$

Die wahre Lösung ist, wie man leicht sieht $u(x) = 1 - x^2$.

In schwacher Form läßt sie sich als Variationsaufgabe formulieren mit:

$$V := H_0^1(-1, +1)$$

$$a(u, v) := \int_{-1}^{+1} u' v' \, dx \quad ; \quad f(v) := 2 \int_{-1}^{+1} v \, dx$$

$$\text{gesucht: } u \in V : \quad v \in V : \quad a(u, v) = f(v)$$

Wir wollen nun eine Ersatzaufgabe suchen, die möglichst nur noch in der linken Intervallhälfte (also in $(-1, 0)$) gestellt ist, aber dort die gleiche Lösung hat.

Da die Funktionen aus V noch stetig sein müssen, ist dies nicht streng möglich. Eine Zerlegung in $V = V_1 + V_2$ in nicht überlappende Teilräume V_1, V_2 läßt sich nicht finden und wir müssen uns deshalb mit überlappender Kondensation begnügen und zulassen, daß sich die Teilräume in einer schmalen Zone $(-\varepsilon, +\varepsilon)$ überlappen.

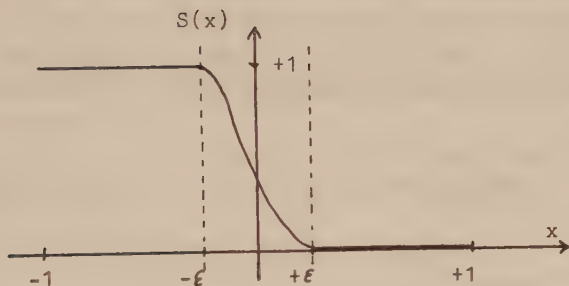
Dazu dient die folgende Hilfskonstruktion:

Sei ein "kleines" $\varepsilon > 0$ vorgegeben und sei $S \in C^\infty(-1, +1)$ eine

glatte Abschneidefunktion mit:

$$S(x) = \begin{cases} +1 & \text{in } -1 \leq x \leq -\epsilon \\ \text{bel.} & \text{in } -\epsilon \leq x \leq +\epsilon \\ 0 & \text{in } +\epsilon \leq x \leq +1 \end{cases}$$

Skizze:



Wir setzen dann:

$$V_1 := \{ v \cdot S(x) \mid v \in V \}$$

$$V_2 := \{ v \cdot (1 - S(x)) \mid v \in V \}$$

Der lineare Operator $L: V_1 \longrightarrow V_2$; $u_1 \longmapsto u_2$ aus Satz (3.1) ist dann definiert über die Lösung des Variationsproblems:

$$\text{gesucht } u_2 \in V_2 : \forall v \in V_2 : a(u_2, v) = a(u_1, v)$$

Wiederum als Differentialgleichung geschrieben entspricht dies der Aufgabe:

$$\begin{aligned} u_2'' &= u_1'' & \text{in } -\epsilon \leq x \leq \epsilon \\ u_2'' &= 0 & \text{in } \epsilon \leq x \leq 1 \\ u_2 &= 0 & \text{in } -1 \leq x \leq -\epsilon \end{aligned}$$

also mit noch unbekannten Koeffizienten a, b, c, d folgt:

$$u_2(x) = \begin{cases} u_1(x) + a + bx & \text{in } -\epsilon \leq x \leq \epsilon \\ c + dx & \text{in } \epsilon \leq x \leq 1 \end{cases}$$

Zur Bestimmung dieser Koeffizienten stehen dabei folgende 4 Bedingungen zur Verfügung:

$$\begin{aligned} u_2(-\varepsilon) &= 0 && \text{weil } u_2 \in V_2 \\ u_2(+1) &= 0 && \text{" " " " } \\ u_2(\varepsilon^-) &= u_2(\varepsilon^+) && \text{weil } u_2 \in C^2 \\ u_2'(\varepsilon^-) &= u_2'(\varepsilon^+) && \text{" " " " } \end{aligned}$$

Insgesamt ergibt sich für den Operator L damit folgende Gestalt:

$$u \in V_1, \quad \text{es sei } u_\varepsilon := u(-\varepsilon)$$

$$Lu(x) = \begin{cases} u(x) + (x-1) \cdot u_\varepsilon / (1+\varepsilon) & \text{in } -\varepsilon \leq x \leq \varepsilon \\ (x-1) \cdot u_\varepsilon / (1+\varepsilon) & \text{in } \varepsilon \leq x \leq 1 \end{cases}$$

Die Bilinearform a^* ergibt sich damit nach partieller Integration, Berücksichtigung von $v \in V_1$, also $v(\varepsilon)=0$, und der Abkürzung $v(-\varepsilon) := v_\varepsilon$ zu:

$$\begin{aligned} a^*(u, v) &= a(u - Lu, v) = \dots = \\ &= \int_{-1}^{-\varepsilon} u' v' dx + u_\varepsilon \cdot v_\varepsilon / (1+\varepsilon) \end{aligned}$$

Die Funktion $u_\varepsilon \in V_2$ findet man auf analogem Wege zu:

$$u_\varepsilon = -x^2 + (1-\varepsilon)x + \varepsilon \quad \text{in } -\varepsilon \leq x \leq +1$$

und damit lautet nun die gesuchte Ersatzaufgabe folgendermaßen:

gesucht $u \in V_1$: $\forall v \in V_1$:

$$\int_{-1}^{-\varepsilon} u' v' dx + u_\varepsilon \cdot v_\varepsilon / (1+\varepsilon) = 2 \cdot \int_{-1}^{+\varepsilon} v dx - \int_{-\varepsilon}^{+\varepsilon} (-2x + (1-\varepsilon)) v' dx$$

und wie man leicht verifizieren kann, ist die Lösung der ursprünglichen Aufgabe $u(x)=1-x^2$ tatsächlich eine Lösung dieses Problems.

Man erkennt, daß der Formalismus bereits bei einfachsten Beispielen relativ aufwendig wird und in dieser Form wohl eher theoretischen als praktischen Nutzen hat. Bei der Anwendung auf finite Elemente ergibt sich daraus genau wieder die statische Kondensation. Der Operator L entspricht der Matrix $k_{22}k_{21}$ und die Bilinearform a^* wird im wesentlichen durch die kondensierte Matrix K^* repräsentiert.

Auch ist es in der diskreten Anwendung immer möglich eine nicht überlappende Zerlegung des Grundraumes anzugeben.

4) VERTRÄGLICHKEIT

Die grundsätzliche Idee dieser Arbeit ist es, anstelle des endlich dimensionalen Teilraumes V_h zu einem Raum mit (abzählbar) unendlich vielen Dimensionen überzugehen und so die Approximationseigenschaften zu verbessern.

Das lineare Gleichungssystem $K U = F$ erhält damit natürlich ebenfalls unendliche Dimension. Der einzige Weg solche unendlichen Systeme tatsächlich aufzustellen und zu bearbeiten, besteht wohl in einer wie auch immer gearteten Rekursionsvorschrift.

Wir benötigen also F.E.-Zerlegungen, bestehend aus (unendlich vielen) einzelnen Substrukturen, deren jeweilige Steifigkeitsmatrizen bzw. Lastvektoren in möglichst einfacher Weise rekursiv auseinanderhervorgehen.

Dazu wollen wir uns in diesem Abschnitt mit der Frage beschäftigen, was mit der zugehörigen Steifigkeitsmatrix und dem Lastvektor passiert, wenn man die komplette Substruktur einer affinen Transformation unterzieht. Genauer gesagt interessieren uns natürlich nur solche Transformationen, bei denen möglichst wenig passiert.

In §5 werden wir dann die Bedingungen diskutieren, die an eine F.E.-Zerlegung zu stellen sind, damit man die transformierte Substruktur mit der originalen koppeln kann. Durch Wiederholung des Prozesses kann man auf diese Weise ein Blocktridiagonal-System erzeugen, das dann in §6 durch rekursive statische Kondensation aufgelöst wird. §7 untersucht schließlich die Konvergenzeigenschaften der entstehenden Rekursionsvorschrift.

Wir wollen das Problem zunächst abstrakt im HILBERT-Raum studieren. Dazu folgende Definition:

(4.1) DEFINITION:

$V(\Omega)$ sei HILBERT-Raum von Funktionen auf Ω .

$D, D' \subset V$, $a: \forall x \in V \longrightarrow \mathbb{R}$ Bilinearform und

$\mathcal{T}: D \longrightarrow D'$ Isomorphismus

a heißt mit $\mathcal{T}$ verträglich zum Faktor $\gamma: \langle == \rangle$

$$\forall u, v \in D: a(\mathcal{T}u, \mathcal{T}v) = \gamma a(u, v)$$

bei $\gamma = 1$ heißt a invariant unter $\mathcal{T}$.

Ganz analog zur Verträglichkeit der Bilinearform a , soll auch für das Funktional $f(v)$ der Variationsaufgabe, aus dem ja die rechte Seite des Gleichungssystems entsteht, eine Verträglichkeit (zum selben γ !) definiert werden:

f verträglich mit $\mathcal{T}$ (zum Faktor γ) $\langle == \rangle$

$$\forall v \in D: f(\mathcal{T}v) = \gamma f(v)$$

Insbesondere interessieren wir uns dabei für Isomorphismen $\mathcal{T}$, die von affinen, bijektiven Abbildungen $\hat{T}$ des $\mathbb{R}^n$ in sich herrühren:

Sei: $\hat{T}: \mathbb{R}^n \longrightarrow \mathbb{R}^n$

$x \longmapsto Tx + b$; mit $T \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$, T invertierbar

und $\mathcal{T}: D \longrightarrow D'$

$u \longmapsto u \circ \hat{T}^{-1} = u(T^{-1}(x-b))$

der davon induzierte Isomorphismus.

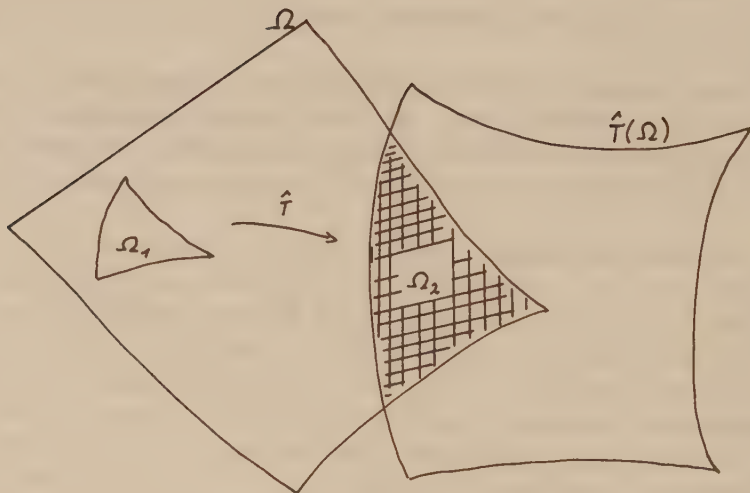
D und D' müssen dabei so gewählt werden, daß $\mathcal{T}$ "innerhalb" von $V(\Omega)$ bleibt. Dies erreicht man beispielsweise durch folgende Konstruktion:

$$\begin{aligned} \text{Sei } \Omega_1 &:= \hat{T}^{-1}(\hat{T}(\Omega) \cap \Omega) \neq \emptyset \\ \Omega_2 &:= \hat{T}(\Omega_1) \end{aligned}$$

$$D := \{ v \in V(\Omega) \mid \text{supp}(v) \subset \Omega_1 \}$$

$$D' := \{ v \in V(\Omega) \mid \text{supp}(v) \subset \Omega_2 \}$$

Skizze:



Vereinfachend wollen wir auch von Verträglichkeit (bzw. Invarianz) von a mit $\hat{T}$ reden, wenn a mit dem zugehörigen $\mathcal{T}$ verträglich (invariant) ist.

Der Bereich der prinzipiell möglichen Werte für γ ist dabei eng verknüpft mit den Konstanten C und M der Bilinearform (siehe Fußnote Seite 1) und den Normen von $\mathcal{T}$ und $\mathcal{T}^{-1}$.

Es gilt:

(4.2) LEMMA:

$$C/M \sup_{u \neq 0} \frac{\|\mathcal{T}u\|^2}{\|u\|^2} = |\gamma| = M/C \inf_{u \neq 0} \frac{\|\mathcal{T}u\|^2}{\|u\|^2}$$

BEWEIS:

$$\begin{aligned} \text{Es ist: } |a(\mathcal{T}u, \mathcal{T}u)| &\geq C \|\mathcal{T}u\|^2 \\ &= \\ | \gamma / a(u, u) | &\leq |\gamma| M \|u\|^2 \quad \text{für alle } u \in D \end{aligned}$$

also $|\gamma| \geq C/M (\|\mathcal{T}u\|/\|u\|)^2$ f. alle u
damit ist die linke Ungleichung bewiesen. Die rechte folgt ganz analog aus:

$$\begin{aligned} |a(\mathcal{T}u, \mathcal{T}u)| &\leq M \|\mathcal{T}u\|^2 \\ &= \\ | \gamma / a(u, u) | &\geq |\gamma| C \|u\|^2 \quad \text{f. alle } u \end{aligned}$$

q.e.d.

Für definite Bilinearformen (d.h. $a(u, u) > 0$ für alle $u \neq 0$ oder $a(u, u) < 0$ für alle $u \neq 0$) sieht man sofort, daß der Verträglichkeitsfaktor positiv sein muß. Der Fall $\gamma = 0$ wird durch die geforderte Elliptizität ausgeschlossen.

Der nun folgende Satz zeigt, wie sich die Steifigkeitsmatrix ändert, wenn man das Grundgebiet in verträglicher Weise transformiert. Wenn man so will ist das der zentrale Satz, auf dem die ganze Arbeit aufbaut.

(4.3) SATZ:

Sei $Z := ((E_i, P_i, \Sigma_i)) \ i=1 \dots k$ F.E.-Zerlegung eines Grundgebietes Ω_0 in affin äquivalente Elemente. K die Steifigkeitsmatrix und F der Lastvektor dazu.

$\hat{T}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ affine bijektive Transformation, die mit der zugrunde gelegten Bilinearform a und dem Funktional f zum Faktor γ verträglich ist.

$Z' := ((E'_i, P'_i, \Sigma'_i)) \ i=1 \dots k$ sei die von $\hat{T}$ erzeugte affin äquivalente F.E.-Zerlegung von $\hat{T}(\Omega_0)$. K^*, F^* Steifigkeitsmatrix bzw. Lastvektor dazu.

Dann gilt:

$$K^* = \gamma K \quad \text{und} \quad F^* = \gamma F$$

BEWEIS:

Die Koeffizienten k_{ij} der Matrix K haben die Form:

$$k_{ij} = a(\psi_i, \psi_j)$$

wobei die ψ_i, ψ_j Basisfunktionen der F.E.-Zerlegung sind.

Da die Zerlegung als affin äquivalent vorausgesetzt ist, gilt wegen der Verträglichkeit:

$$k_{ij}^* = a(\psi_i \circ \hat{T}^{-1}, \psi_j \circ \hat{T}^{-1}) = \gamma a(\psi_i, \psi_j) = \gamma k_{ij}.$$

Analog für die Komponenten F_i^* von F^* :

$$F_i^* = f(\psi_i \circ \hat{T}^{-1}) = \gamma f(\psi_i) = \gamma F_i$$

q.e.d.

Damit das ganze nicht leere Theorie bleibt, sollen nun ein paar konkrete Bilinearformen aus praktisch relevanten Fällen auf Verträglichkeit hin untersucht werden.

Zunächst ein hinreichendes und notwendiges Kriterium für allgemeine elliptische Probleme 2. Ordnung im $\mathbb{R}^n$.

(4.4) SATZ:

Sei $\hat{T}: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x \mapsto Tx+b$, ($T \in \mathbb{R}^{n \times n}$ invertierbar, $b \in \mathbb{R}^n$)

$V := H^1(\Omega)$

$$a(u,v) := \int_{\Omega} \sum_{i,j=1}^n a_{ij}(x) \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx; \quad A(x) := (a_{ij}(x))_{i,j=1 \dots n}$$

Dann gilt:

a ist mit $\hat{T}$ verträglich $\Leftrightarrow$

$$\exists \gamma: \text{für fast alle } x \in \Omega : T^T A(T^{-1}(x-b)) T = \gamma |\det(T)| A(x)$$

BEWEIS:

Seien $u, v \in V$ so, daß auch $u \circ \hat{T}^{-1}$, $v \circ \hat{T}^{-1} \in V$

Die Definition von Verträglichkeit lautet:

$$a(u \circ \hat{T}^{-1}, v \circ \hat{T}^{-1}) = \gamma a(u, v)$$

$$\Leftrightarrow \int_{\Omega} \sum_{i,j} a_{ij}(x) \frac{\partial (u \circ \hat{T}^{-1})}{\partial x_i} \frac{\partial (v \circ \hat{T}^{-1})}{\partial x_j} dx = \gamma \int_{\Omega} \sum_{i,j} a_{ij}(x) \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx$$

oder in Matrixschreibweise ausgedrückt:

$$\Leftrightarrow \int_{\Omega} (\nabla(u \circ \hat{T}^{-1}))^T A(x) \nabla(v \circ \hat{T}^{-1}) dx = \gamma \int_{\Omega} \nabla u^T A(x) \nabla v dx$$

mit der Kettenregel folgt:

$$\Leftrightarrow \int_{\Omega} \nabla u^T (\hat{T}^{-1})^T T^{-T} A(x) T^{-1} \nabla v dx = \gamma \int_{\Omega} \nabla u^T A(x) \nabla v dx$$

Auf die linke Seite dieser Gleichung wird nun eine Integraltransformation mit $\hat{T}$ angewendet. Das Integrationsgebiet Ω ändert sich dabei formal nicht, da nur der Träger des Integranden in Ω verschoben wird.

$$\Leftrightarrow \int_{\Omega} \nabla u^T T^T A(T^{-1}(x)) T |(1/\det(T))| \nabla v dx = \gamma \int_{\Omega} \nabla u^T A(x) \nabla v dx$$

da dies für alle u, v gelten muß, folgt:

$$T^T A(T^{-1}(x)) T = \gamma |\det(T)| A(x)$$

q.e.d.

(4.5) BEISPIELE:

Die Bilinearform zum n-dimensionalen LAPLACE-Operator

$$-\Delta u = -\left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \dots + \frac{\partial^2 u}{\partial x_n^2} \right)$$

ist bekanntlich gegeben durch

$$a(u, v) = \int_{\Omega} \text{grad}(u) \cdot \text{grad}(v) \, dx$$

Mit dem vorherigen Satz folgt sofort: ($A = \text{Einheitsmatrix}$)

- (i) a ist verträglich mit zentrischen Streckungen $\hat{T} = tI$
($I = \text{Einheitsmatrix}$) zum Faktor $\gamma = t^{n-2}$
Insbesondere also invariant in 2 Dimensionen ($n=2$) !!

- (ii) a ist invariant bei allen Translationen $\hat{T}x = x+b$
($b \in \mathbb{R}^n$)

- (iii) a ist invariant unter allen Orthogonaltransformationen
 $T=Q$; mit $Q^T Q = I$

Die Bilinearform zum Operator der HELMHOLTZ-Gleichung $\Delta u + k^2 u$ ist gegeben durch:

$$a(u, v) = \int_{\Omega} -\text{grad}(u) \cdot \text{grad}(v) + k^2 uv \, dx$$

Der Anteil $k^2 uv$ im Integral führt dazu, daß diese Bilinearform nicht verträglich ist mit zentrischen Streckungen. Nur unter Transformationen mit $|\det T| = 1$ und Translationen bleibt sie invariant. Damit ist das in dieser Arbeit vorgeschlagene Verfahren zur Behandlung von Singularitäten aber leider hier nicht anwendbar.

Ein Beispiel für einen Operator 4. Ordnung ist der sog. Bipotentialoperator, der unter anderem das elastische Verhalten dünner

Platten beschreibt:

$$\nabla^4 u = \Delta \Delta u = \frac{\partial^4 u}{\partial x_1^4} + 2 \frac{\partial^4 u}{\partial x_1^2 \partial x_2^2} + \frac{\partial^4 u}{\partial x_2^4}$$

Er führt auf folgende Bilinearform: (siehe z.B: CIARLET(1978))

$$a(u, v) = \int_{\Omega} \Delta u \Delta v \, dx \quad ; u, v \in H^2(\Omega)$$

Auch diese Bilinearform hat günstige Verträglichkeitseigenschaften:

(4.6) SATZ:

Die Bilinearform des Bipotentialoperators ist invariant unter Translationen und verträglich mit zentrischen Streckungen.

BEWEIS:

Sei $t > 0$, $b \in \mathbb{R}^n$, $\hat{T}: \mathbb{R}^n \longrightarrow \mathbb{R}^n$, $x \longmapsto Tx + b$; $T = \text{diag}(t)$
 $u, v \in H^2(\Omega)$ derart, daß auch $u \circ \hat{T}^{-1}$, $v \circ \hat{T}^{-1} \in H^2(\Omega)$

$$\begin{aligned} a(u \circ \hat{T}^{-1}, v \circ \hat{T}^{-1}) &= \int_{\Omega} \Delta(u \circ \hat{T}^{-1}) \Delta(v \circ \hat{T}^{-1}) \, dx = \\ &= \int_{\Omega} t^4 (\Delta u) \circ \hat{T}^{-1} (\Delta v) \circ \hat{T}^{-1} \, dx = \end{aligned}$$

durch Integraltransformation mit $\hat{T}$ ergibt sich:

$$= t^4 / |\det(T)| \, a(u, v)$$

dabei bleibt das Integrationsgebiet Ω formal unverändert, da sich nur der Träger des Integranden in Ω verschiebt.

Mit $|\det(T)| = t^n$ folgt sofort die Behauptung. Der Verträglichkeitsfaktor ist $\gamma = t^{4-n}$.

q.e.d.

Ein sehr wichtiges Anwendungsgebiet für die Methode der Finiten Elemente ist der Bereich der Elastizitätstheorie. In Anlehnung an CIARLET (1978) (Ch. 1, §1.2) lassen sich die beschreibenden Gleichungen hier folgendermaßen als Variationsaufgabe der Form (1.0) formulieren:

Sei $V := \underline{u} = (u_1, u_2, u_3) \in H^1(\Omega)^3$; $u_i = 0$ auf Γ_0 , $i=1,2,3$
 dabei ist Γ_0 der Teil des Randes von Ω , auf dem der betrachtete Körper fest eingespannt ist (wesentliche Randbedingung).

mit den Tensoren

$$\epsilon_{ij}(\underline{u}) = 1/2 (\partial_i u_j + \partial_j u_i) \quad i, j=1,2,3 \quad \text{"Dehnung"}$$

und

$$\sigma_{ij}(\underline{u}) = \lambda \left(\sum_k \epsilon_{kk}(\underline{u}) \right) \delta_{ij} + 2\mu \epsilon_{ij}(\underline{u}) \quad i, j=1,2,3 \quad \text{"Spannung"}$$

ist die Bilinearform a gegeben durch:

$$a(u, v) := \int_{\Omega} \sum_{i,j=1}^3 \sigma_{ij}(\underline{u}) \epsilon_{ij}(\underline{v}) \, dx$$

dabei ist δ_{ij} das KRONECKER-Symbol und λ, μ sind Materialkonstanten (die sog. LAME-Konstanten).

Zusammen mit einer rechten Seite $f(\underline{v})$, die eventuell einwirkende Volumen- und Oberflächenkräfte repräsentiert, beschreibt dann das Variationsproblem (1.0) die elastischen Verschiebungen eines 3-dimensionalen isotropen Körpers unter der Einwirkung von Kräften.

(4.7) SATZ:

Die Bilinearform der isotropen 3d- Elastizitätstheorie ist invariant unter Translationen und verträglich mit zentrischen Streckungen. Der Verträglichkeitsfaktor γ ist dabei gleich dem Streckungsfaktor.

BEWEIS:

Sei $t > 0$, $b \in \mathbb{R}^3$, $\hat{T}: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, $x \longmapsto Tx + b$, $T = \text{diag}(t)$

Die Umkehrung von $\hat{T}$ ist:

$$\hat{T}^{-1}(x) = 1/t \ x - b' \ , \ \text{mit } b' = 1/t \ b \ .$$

Es gilt:

$$\epsilon_{ij}(\underline{y} \circ \hat{T}^{-1})(x) = 1/2 \left(\frac{\partial}{\partial x_i} (u_j(1/t \ x - b')) + \frac{\partial}{\partial x_j} (u_i(1/t \ x - b')) \right) =$$

mit der Kettenregel folgt

$$\begin{aligned} &= 1/2 \cdot \left(\frac{\partial}{\partial x_i} u_j(\hat{T}^{-1}x) \frac{1}{t} + \frac{\partial}{\partial x_j} u_i(\hat{T}^{-1}x) \frac{1}{t} \right) = \\ &= \frac{1}{t} \epsilon_{ij}(\underline{y}) \circ \hat{T}^{-1}(x) \end{aligned}$$

analog gilt für die Spannungen:

$$\begin{aligned} \sigma_{ij}(\underline{y} \circ \hat{T}^{-1})(x) &= \lambda (\sum_{kk} \epsilon_{kk}(\underline{y} \circ \hat{T}^{-1})) \delta_{ij} + 2\mu \epsilon_{ij}(\underline{y} \circ \hat{T}^{-1}) = \\ &= (\lambda (\frac{1}{t} \sum_{kk} \epsilon_{kk}(\underline{y})) \delta_{ij} + \frac{2}{t} \mu \epsilon_{ij}(\underline{y})) (\hat{T}^{-1}x) = \\ &= \frac{1}{t} \sigma_{ij}(\underline{y}) \circ \hat{T}^{-1}(x) \end{aligned}$$

setzt man dies in die Bilinearform ein, so folgt:

$$\begin{aligned} a(\underline{y} \circ \hat{T}^{-1}, \underline{y} \circ \hat{T}^{-1}) &= \int_{\Omega} \sum_{i,j} \epsilon_{ij}(\underline{y} \circ \hat{T}^{-1}) \sigma_{ij}(\underline{y} \circ \hat{T}^{-1}) \, dx = \\ &= \frac{1}{t^2} \int_{\Omega} \sum_{i,j} \epsilon_{ij}(\underline{y}) \circ \hat{T}^{-1} \sigma_{ij}(\underline{y}) \circ \hat{T}^{-1} \, dx = \end{aligned}$$

Durch eine Integraltransformation mit $\hat{T}$ folgt (das Integrationsgebiet bleibt dabei wieder formal gleich, da sich nur der Träger des Integranden verschiebt):

$$\begin{aligned} &= t \int_{\Omega} \sum_{i,j} \epsilon_{ij}(\underline{y}) \sigma_{ij}(\underline{y}) \, dx = \\ &= t \, a(u, v) \end{aligned}$$

q.e.d.

(4.8) KOROLLAR:

Die Bilinearform der ebenen Elastizitätstheorie ist invariant unter zentrischen Streckungen und unter Parallelverschiebungen.

BEWEIS:

Setzt man im Satz (4.5) $u_3=0$ und betrachtet die affine Transformation nur im $\mathbb{R}^2$, so ist bei der Integraltransformation $|\det(T)| = t^2$ anstelle von t^3 und die Behauptung folgt unmittelbar.

q.e.d.

Zusammen mit Satz (4.3) ergibt sich so die vielleicht zunächst etwas verblüffende Tatsache, daß die Steifigkeitsmatrix einer 2D-Struktur völlig unverändert bleibt, wenn man diese maßstabsgetreu vergrößert oder verkleinert, daß aber bei einer 3D-Struktur die Matrix mit dem Streckungsfaktor zu multiplizieren ist.

Man kann sich dies an einem kleinen eindimensionalen Beispiel plausibel machen:

Ein einfacher Stab, der Länge l mit Querschnittsfläche A und E-Modul E wird durch eine Kraft F in Längsrichtung gezogen.

Die Dehnung x läßt sich dann leicht aus der bekannten Formel

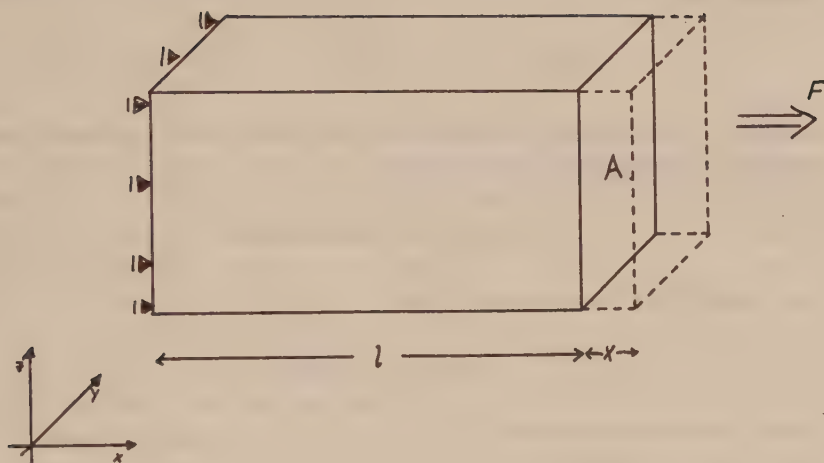
$$F = K x$$

berechnen. Die Federkonstante K ergibt sich dabei aus

$$K = EA/l$$

Dies ist gewissermaßen ein 1-dimensionales Analogon zu dem linearen Gleichungssystem bei der F.E.-Methode. Die Federkonstante K entspricht dabei der Steifigkeitsmatrix.

Skizze:



bei zentrischer Streckung in 3 Dimensionen hat man nun:

$$\begin{array}{lcl} A & \longrightarrow & t^2 A \quad (\text{Fläche}) \\ l & \longrightarrow & t l \end{array}$$

also $K \longrightarrow t K$

bei Betrachtung in 2 Dimensionen gilt aber:

$$\begin{array}{lcl} A & \longrightarrow & t A \quad (\text{Linie}) \\ l & \longrightarrow & t l \end{array}$$

also $K \longrightarrow K$

Wir wollen nun noch kurz die Verträglichkeit der rechten Seite, d.h. des Funktionals $f(v)$ etwas näher beleuchten. Die homogene Aufgabe, also $f(v)=0$ ist trivialerweise stets mit allen Transformationen verträglich. Die Art von Verträglichkeit, auf die es im weiteren besonders ankommen wird, ist die Verträglichkeit mit zentrischer Streckung.

Wenn wir also davon ausgehen, daß das Funktional $f(v)$ die übliche Gestalt hat:

$$f(v) = \int_{\Omega} \bar{f} v \, dx$$

mit einer vorgegeben Funktion $\bar{f} \in L_2(\Omega)$, dann ist die Frage, welche Gestalt $\bar{f}$ haben muß, damit gilt:

$$\bar{f}(T(x)) = \gamma / \det(T) \bar{f}(x)$$

Für zentrische Streckungen also:

$$\bar{f}(tx) = \gamma t^n \bar{f}(x)$$

Beim 2d LAPLACE-Operator war z.B. $\gamma = 1$ (Invarianz) und es ist also von f zu fordern:

$$\bar{f}(tx) = t^2 \bar{f}(x)$$

d.h. für f ist eine Funktion folgender Gestalt zulässig:

$$\bar{f}(r, \phi) = r^2 g(\phi)$$

mit Polarkoordinaten r, ϕ bezogen auf das Streckungszentrum und einer beliebigen Funktion $g(\phi)$.

Beim 3d LAPLACE-Operator war $\gamma = t$ also ist vom f zu fordern:

$$\bar{f}(r, \phi) = r^4 g(\phi)$$

5) REKURSIVE F.E.-ZERLEGUNG

Im vorigen Abschnitt hatten wir das Verhalten einer Steifigkeitsmatrix bei Affintransformation des Grundgebietes untersucht. Unser Ziel ist es nun, mehrere Substrukturen miteinander zu koppeln, die alle durch wiederholtes Anwenden ein und der selben Transformation auseinander hervorgehen.

In diesem Abschnitt soll diskutiert werden, welcher Art Strukturen auf diese Weise erzeugt werden können, welchen Bedingungen das Grundelement genügen muß und welche Matrizen dabei entstehen.

Sei also $\hat{T} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ bijektive affine Transformation.

Wir wollen davon ausgehen, daß das Grundgebiet Ω in folgender Weise zerlegbar sei.

$$\Omega = \Omega^* \cup \bar{R}_0 \cup \bar{R}_1 \cup \bar{R}_2 \cup \dots \cup \bar{R}_N \quad (5.1)$$

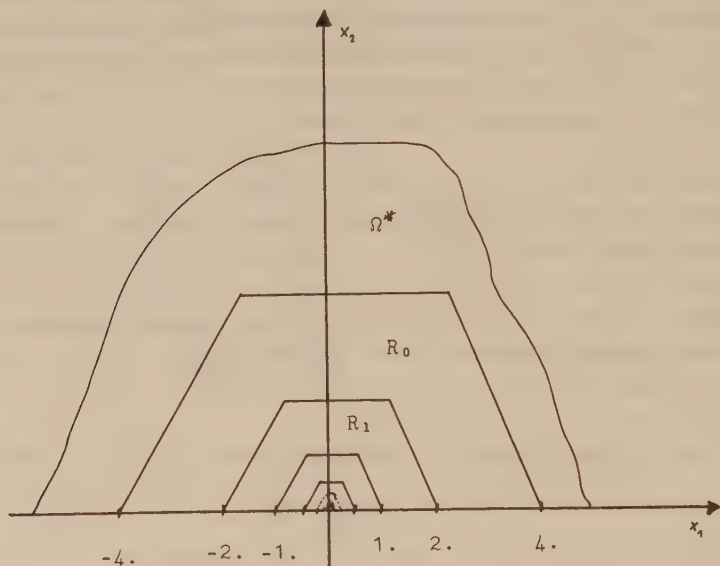
mit $\bar{R}_i = \hat{T} \bar{R}_{i-1} \quad i=1,2,\dots,N ; \quad R_i \text{ offen}$

$$\text{und für } i \neq j: \quad \bar{R}_i \cap \bar{R}_j = \begin{cases} \text{gemeinsames Randstück für } |i-j|=1 \\ \emptyset & \text{für } |i-j| > 1 \end{cases} \quad (5.2)$$

Ω^* enthält dabei das Restgebiet, das sich nicht in dieses Schema pressen läßt.

Eine schematische Skizze für den Spezialfall einer zentrischen Stauchung um dem Faktor $1/2$ mag dies verdeutlichen:

Skizze:



Die geeignete Wahl von $\hat{T}$ und R_0 für eine vorgegebene Anwendung kann unter Umständen sehr schwierig oder gar unmöglich sein. Für ein fest vorgegebenes $\hat{T}$ läßt sich eine Zerlegung mit obigen Eigenschaften aber stets konstruieren, wenn nur $\hat{T}$ "in Ω bleibt", d.h. $\hat{T}\Omega \subset \Omega$.

Ausgehend von einem Ω_0 mit der Eigenschaft $\hat{T}\Omega_0 \cap \Omega_0 \neq \emptyset$ kann man aus der Folge der Mengen

$$\Omega_0, \hat{T}\Omega_0, \hat{T}^2\Omega_0, \hat{T}^3\Omega_0, \dots$$

geeignete Mengen R_i finden durch eine einfache Mengendifferenz.
Man muß dabei zwei Fälle unterscheiden:

1. Fall: $\Omega_0 \subset \hat{T} \Omega_0$:

$$\text{setze: } R_0 := \overbrace{(\hat{T} \Omega_0 \setminus \Omega_0)}^o; \quad \bar{R}_i = \hat{T} \bar{R}_{i-1} \quad i=1, 2, \dots, N$$

$$\Omega^* := \Omega \setminus (\bar{R}_0 \cup \bar{R}_1 \cup \bar{R}_2 \cup \dots \bar{R}_N)$$

2. Fall: sonst

$$\text{setze: } R_0 := \overbrace{(\Omega_0 \setminus \hat{T} \Omega_0)}^o; \quad \bar{R}_i = \hat{T} \bar{R}_{i-1} \quad , \quad i=1, 2, \dots, N$$

$$\Omega^* := \Omega \setminus (\bar{R}_0 \cup \bar{R}_1 \cup \bar{R}_2 \cup \dots \bar{R}_N)$$

Der "Grundring" R_0 sei polygonal berandet und in affin äquivalente Elemente zerlegt. Durch die Transformation $\hat{T}$ werden dann auch alle nachfolgenden R_i in analoger Weise zerlegt. Man erzeugt so eine Folge von affin äquivalenten Substrukturen $R_0, R_1, R_2, \dots$

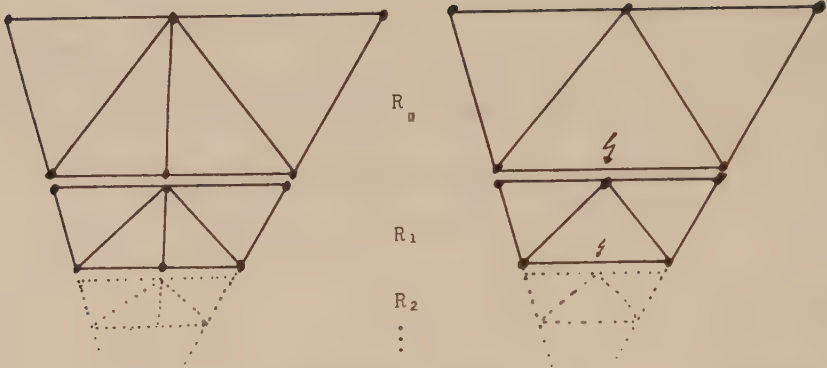
Wir wollen fordern, daß diese Substrukturen zusammen ebenfalls wieder eine zulässige F.E.-Zerlegung bilden, also auch die Vereinigung noch die Konnektivitätsbedingung (1.7) erfüllt.

Dazu müssen aufeinanderfolgende Gebiete $\bar{R}_{i-1}, \bar{R}_i \quad i=1, 2, \dots, N$ längs des gemeinsamen Randstückes an den gleichen Stellen Knoten haben und gleiche Ansatzfunktionen.

Damit das möglich ist, muß die Zerlegung von R_0 so sein, daß das Randstück $\hat{T} \bar{R}_0 \cap \bar{R}_0$ mit seinem Urbild unter $\hat{T}$, also $\hat{T}^{-1}(\hat{T} \bar{R}_0 \cap \bar{R}_0)$ affin äquivalent ist.

Eine Skizze soll das wieder verdeutlichen:

Skizze:



erlaubt

nicht erlaubt

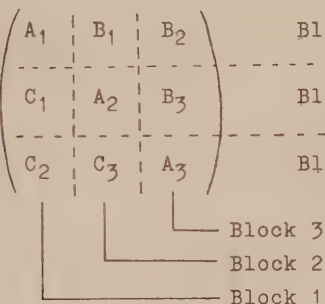
Auf diese Weise wird also aus einer einzigen Substruktur durch wiederholte affine Transformation eine F.E.-Zerlegung von $\bigcup_{i=0}^N R_i$ generiert. Diese kann nun als Macro-Element aufgefasst werden und ganz normal an eine (kompatible) Zerlegung von Ω^* angekoppelt werden.

Letzten Endes interessiert uns speziell der Fall $N \rightarrow \infty$. Zunächst soll aber erst noch der Aufbau des so erzeugten linearen Gleichungssystems zu $\bigcup_{i=0}^N R_i$ näher untersucht werden.

Sei S_0 die Steifigkeitsmatrix und F_0 der Lastvektor zu R_0 . Die zugrundeliegende Bilinearform und die rechte Seite seien verträglich mit $\hat{T}$ zum Faktor γ . Dann sind die entsprechenden Größen S_i , F_i von R_i offenbar gegeben durch:

$$\begin{aligned} S_i &= \gamma^i S_0 \\ F_i &= \gamma^i F_0 \end{aligned}$$

Damit man diese Teilmatrizen zur Gesamtsteifigkeitsmatrix zusammensetzen kann, muß man sie folgendermaßen partitionieren (eventuell nach geeigneter Umnummerierung der Knoten):

$$S_i = \gamma^i \begin{pmatrix} A_1 & B_1 & B_2 \\ C_1 & A_2 & B_3 \\ C_2 & C_3 & A_3 \end{pmatrix} \begin{matrix} \text{Block 1} \\ \text{Block 2} \\ \text{Block 3} \end{matrix}$$


Analog für die Lastvektoren F_i :

$$F_i = \gamma^i \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} \begin{matrix} \text{Block 1} \\ \text{Block 2} \\ \text{Block 3} \end{matrix}$$

Dabei enthalten die einzelnen Blöcke jeweils:

Block 1: Knotenvariable, die auch im Vorgänger vorkommen, also auf dem gemeinsamen Rand von $\bar{R}_{i-1}$ und $\bar{R}_i$

Block 2: freie Knoten in R_i , die weder mit R_{i+1} noch mit R_{i-1} koppeln. Dieser Block kann auch fehlen. Dann hat man es nur mit einem 2x2-Blocksystem zu tun.

Block 3: Knoten, die mit dem Nachfolger koppeln, also auf dem gemeinsamen Rand von $\bar{R}_i$ mit $\bar{R}_{i+1}$ liegen.

Beim Zusammensetzen der einzelnen Teilmatrizen S_i ist nun zu beachten, daß sich jeweils die Beiträge aus Block 3 von R_i mit Block 1 von R_{i+1} überlagern (aufaddieren). Insgesamt erhält man so für N hintereinandergeschachtelte Substrukturen die Gesamtsteifigkeitsmatrix zu : (N=3)

$$\begin{array}{l}
 R_0 \dots \dots \dots \begin{array}{|ccc|} \hline A_1 & B_1 & B_2 \\ \hline C_1 & A_2 & B_3 \\ \hline C_2 & C_3 & A_3 + \gamma A_1 \quad \gamma B_1 \quad \gamma B_2 \\ \hline \end{array} \\
 R_1 \dots \dots \dots \begin{array}{|cc|} \hline \gamma C_1 \quad \gamma A_2 \quad \gamma B_3 \\ \hline \gamma C_2 \quad \gamma C_3 \quad \gamma A_3 + \gamma^2 A_1 \quad \gamma^2 B_1 \quad \gamma^2 B_2 \\ \hline \end{array} \\
 R_2 \dots \dots \dots \begin{array}{|ccc|} \hline \gamma^2 C_1 \quad \gamma^2 A_2 \quad \gamma^2 B_3 \\ \hline \gamma^2 C_2 \quad \gamma^2 C_3 \quad \gamma^2 A_3 + \gamma^3 A_1 \quad \gamma^3 B_1 \quad \gamma^3 B_2 \\ \hline \end{array} \\
 R_3 \dots \dots \dots \begin{array}{|ccc|} \hline \gamma^3 C_1 \quad \gamma^3 A_2 \quad \gamma^3 B_3 \\ \hline \gamma^3 C_2 \quad \gamma^3 C_3 \quad \gamma^3 A_3 \\ \hline \end{array}
 \end{array} \quad (5.3)$$

Ganz analog baut sich die rechte Seite des Gleichungssystems aus den F_i auf:

$$\left(\begin{array}{c} f_1 \\ f_2 \\ f_3 + \gamma f_1 \\ \gamma f_2 \\ \gamma f_3 + \gamma^2 f_1 \\ \gamma^2 f_2 \\ \gamma^2 f_3 + \gamma^3 f_1 \\ \gamma^3 f_2 \\ \gamma^3 f_3 \end{array} \right) \quad (5.4)$$

Diese Matrix und rechte Seite werden nun jeweils in 2er Blöcke partitioniert und jeweils die i. Blockzeile mit dem Faktor γ^{2-i} durchmultipliziert. Dabei ändert sich ja die Lösung nicht.

Schema:

$$\begin{array}{l}
 \gamma^0 \cdot \begin{array}{|cc|cc|} \hline \gamma A_1 & \gamma B_1 & rB_2 & 0 \\ \hline \gamma C_1 & \gamma A_2 & rB_3 & 0 \\ \hline \end{array} \\
 \gamma^1 \cdot \begin{array}{|cc|cc|cc|} \hline C_2 & C_3 & A_3 + \gamma A_1 & rD_1 & rB_2 & 0 \\ \hline 0 & 0 & rC_1 & rA_2 & rB_3 & 0 \\ \hline \end{array} \\
 \gamma^2 \cdot \begin{array}{|cc|cc|cc|cc|} \hline & C_2 & C_3 & A_3 + \gamma A_1 & rD_1 & rB_2 & 0 \\ \hline & 0 & 0 & rC_1 & rA_2 & rB_3 & 0 \\ \hline \end{array} \\
 \gamma^3 \cdot \begin{array}{|cc|cc|cc|cc|cc|} \hline & & C_2 & C_3 & A_3 + \gamma A_1 & rD_1 & rB_2 & 0 \\ \hline & & 0 & 0 & rC_1 & rA_2 & rB_3 & 0 \\ \hline \end{array} \\
 \gamma^4 \cdot \begin{array}{|cc|cc|cc|cc|cc|cc|} \hline & & & C_2 & C_3 & A_3 + \gamma A_1 & rD_1 & rB_2 & 0 \\ \hline & & & 0 & 0 & rC_1 & rA_2 & rB_3 & 0 \\ \hline \end{array} \\
 \gamma^5 \cdot \begin{array}{|cc|cc|cc|cc|cc|cc|cc|} \hline & & & & C_2 & C_3 & A_3 & 0 \\ \hline & & & & 0 & 0 & 0 & I \\ \hline \end{array}
 \end{array} \quad (5.5)$$

Auf diese Weise entsteht ein Blocktridiagonalsystem mit (fast) nur gleichen Zeilen.

Die einzelnen Blockmatrizen sollen im weiteren wie folgt abgekürzt werden:

Diagonalblock: $A_\gamma := \begin{pmatrix} A_3 + \gamma A_1 & \gamma B_1 \\ \gamma C_1 & \gamma A_2 \end{pmatrix}$

obere Nebendiagonale: $B := \begin{pmatrix} B_2 & 0 \\ B_3 & 0 \end{pmatrix}$

untere Nebendiagonale: $C := \begin{pmatrix} c_2 & c_3 \\ 0 & 0 \end{pmatrix}$

Der oberste und der unterste Diagonalblock erhält dabei eine Sonderform:

oberster Diagonalblock: $A_o := \begin{pmatrix} \gamma A_1 & \gamma B_1 \\ \gamma C_1 & \gamma A_2 \end{pmatrix}$

unterster Diagonalblock: $A_u := \begin{pmatrix} A_3 & 0 \\ 0 & I \end{pmatrix}$

Dabei haben wir im untersten Block eine entsprechend große Einheitsmatrix I angefügt, damit alle Blöcke formal gleiche Größe haben. Die rechte Seite wird analog mit Nullen aufgefüllt und die zugehörigen Unbekannten im Gleichungssystem haben reine Dummyfunktion.

Der Vektor der rechten Seite, der sich ergibt, wenn man alle Operationen analog ausführt, hat folgende Form:

$$\begin{pmatrix} F_o \\ F \\ F \\ \cdot \\ \cdot \\ \cdot \\ F \\ F_u \end{pmatrix} \quad \text{mit:} \quad F := \begin{pmatrix} f_3 + \gamma f_1 \\ \gamma f_2 \end{pmatrix}$$

$$F_o := \begin{pmatrix} \gamma f_1 \\ \gamma f_2 \end{pmatrix}$$

$$F_u := \begin{pmatrix} f_3 \\ 0 \end{pmatrix} \quad \leftarrow \text{als Dummy-variablen eingefügt}$$

Alles in allem findet man so das folgende Blocktridiagonalsystem:

$$\begin{pmatrix}
 A_0 & \gamma_B & & & \\
 C & A_\gamma & \gamma_B & & \\
 & C & A_\gamma & \gamma_B & \\
 & & C & A_\gamma & \gamma_B \\
 & & & C & A_u
 \end{pmatrix} \cdot \begin{pmatrix} U_0 \\ U_1 \\ U_2 \\ U_3 \\ \vdots \\ U_N \end{pmatrix} = \begin{pmatrix} F \\ F \\ F \\ F \\ \vdots \\ F \end{pmatrix} \quad (5.6)$$

Die Vektoren U_i enthalten dabei jeweils die gesuchten Knotenwerte zu einer Substruktur R_i , die nicht auch im Nachfolger R_{i+1} vorkommen.

Dieses Gleichungssystem entspricht nun dem Macroelement $\bigcup_{i=0}^N R_i$. Es kann als ganz normale Elementsteifigkeitsmatrix, bzw. Elementlastvektor an das entsprechende System des Restgebietes Ω^* angekoppelt werden.

Je nach der Art der Nummerierung wird diese Kopplung "oben", d.h. über U_0 oder "unten", d.h. über U_N vorgenommen.

6) REKURSIONSFORMEL

Für $N \rightarrow \infty$, also bei Verschachtelung von immer mehr Substrukturen, wird natürlich auch das lineare Gleichungssystem unendlich groß.

THATCHER (1975) benützt in seiner Arbeit für die Auflösung dieses Systems ein Verfahren, das dem Mehrschrittverfahren bei gewöhnlichen Differentialgleichungen ähnelt. Er muß dazu allerdings die Invertierbarkeit der Matrix B fordern. Bei seiner speziellen F.E.-Zerlegung aus einfachen Ringen mit TRIM3 Elementen ist dies in der Tat erfüllt. Im allgemeinen ist es aber eher die Ausnahme als die Regel.

Insbesondere immer dann, wenn Block 2 der Partitionierung tatsächlich vorhanden ist, ist B singulär. Auch die von THATCHER (1976) geforderte Symmetrie von B ist nur in Sonderfällen gegeben.

Deshalb wollen wir hier einen anderen Weg einschlagen. Durch rekursive Anwendung der statischen Kondensation (siehe §2) läßt sich für jede N -fache Verschachtelung eine Rekursionsformel herleiten, die nur noch die Invertierbarkeit von A_γ fordert. Da A_γ aber ja Diagonalblock einer Steifigkeitsmatrix ist, ist dies keine sehr einschneidende Bedingung. (vgl. Lemma (2.1))

Prinzipiell gibt es 2 Möglichkeiten die Kondensation durchzuführen:

Fall (a): "von unten nach oben", d.h. wir beginnen mit U_N , kondensieren dann U_{N-1} , ...usw. bis U_0

Fall (b): "von oben nach unten", d.h. zuerst wird U_0 kondensiert, dann U_1 usw.

Welcher Weg sinnvoller ist, hängt davon ab, wo das Tridiagonalsystem an das Gesamtsystem angekoppelt ist.

Bei der Kondensation "von unten nach oben" erhalten wir im 1. Schritt ein verkleinertes System, das nur noch $U_0, U_1, \dots, U_{N-1}$ enthält:

$$\begin{pmatrix} A_0 & \gamma B & & & \\ C & A_1 & \gamma B & & \\ & C & A_2 & \gamma B & \\ & & C & A_3 & \gamma B \\ & & & \ddots & \ddots \\ & & & & C & K_1 \end{pmatrix} \cdot \begin{pmatrix} U_0 \\ U_1 \\ U_2 \\ U_3 \\ \vdots \\ \vdots \\ U_{N-1} \end{pmatrix} = \begin{pmatrix} F \\ F \\ F \\ F \\ \vdots \\ \vdots \\ H_1 \end{pmatrix}$$

Dabei sind K_1 und H_1 gegeben durch

$$\begin{aligned} K_1 &= A_1 - \gamma B A_u^{-1} C \\ H_1 &= F - \gamma B A_u^{-1} F \end{aligned}$$

und wenn U_{N-1} bekannt ist, dann finden wir U_N aus

$$U_N = A_u^{-1} F - X_0 U_{N-1}$$

$$\text{mit } X_0 = A_u^{-1} C$$

Im nächsten Schritt wird dann U_{N-1} kondensiert, es ergeben sich Matrizen K_2 , X_1 und ein Vektor H_2 usw...

Insgesamt erreichen wir folgende Rekursionsformeln, deren Allgemeingültigkeit sich durch vollständige Induktion leicht beweisen läßt:

$$\begin{aligned} \text{Startwerte: } K_0 &:= A_u \\ X_0 &:= K_0^{-1} C \end{aligned} \quad (6.1)$$

$$\begin{aligned} \text{Iteration: } X_i &= K_i^{-1} C \\ K_{i+1} &= A_\gamma - \gamma B X_i \\ H_{i+1} &= F - \gamma B K_i^{-1} H_i \end{aligned} \quad \left. \vphantom{\begin{aligned} X_i &= K_i^{-1} C \\ K_{i+1} &= A_\gamma - \gamma B X_i \\ H_{i+1} &= F - \gamma B K_i^{-1} H_i \end{aligned}} \right\} \text{ für } i=0,1,\dots,N-1 \quad (6.2)$$

Das iterative Verfahren, das HOUDE HAN (1982) angibt, beruht im wesentlichen darauf, diese Rekursionsformeln in dieser Weise auszuführen.

Setzt man noch K_{i+1} in die Formel für X_{i+1} ein, so erhält man folgende einfache Rekursion:

$$X_{i+1} = (A_\gamma - \gamma B X_i)^{-1} C \quad \text{für } i=0,1,2,\dots,N-1 \quad (6.3)$$

Am Ende bleibt das Gleichungssystem übrig,

$$\begin{pmatrix} A_0 & \gamma B \\ C & K_N \end{pmatrix} \begin{pmatrix} U_0 \\ U_1 \end{pmatrix} = \begin{pmatrix} F \\ H_N \end{pmatrix} \quad (6.4)$$

das anschließend noch einmal kondensiert werden könnte - dieser Schritt paßt allerdings formal nicht ganz in die Rekursion, da hier ein anderer Diagonalterm auftritt (A_0 statt A_γ).

Kennt man erst einmal U_0 und U_1 , so ergeben sich die restlichen Unbekannten rekursiv aus:

$$U_i = K_{i-1}^{-1} H_{i-1} - X_{N-i} U_{i-1} \quad i=2,\dots,N \quad (6.5)$$

Für die Kondensation "von oben nach unten" geht man völlig analog vor. Es ergeben sich ähnliche Rekursionsformeln, die im wesentlichen nur C und B vertauschen:

$$\begin{aligned} \text{Startwerte: } K_0 &:= A_0 \\ X_0 &:= A_0^{-1} B \\ H_0 &:= F \end{aligned} \quad (6.6)$$

$$\text{Iteration: } \left. \begin{aligned} X_i &= \gamma K_i^{-1} B \\ K_{i+1} &= A_\gamma - C X_i \\ H_{i+1} &= F - C K_i^{-1} H_i \end{aligned} \right\} \text{ für } i=0,1,\dots,N-1 \quad (6.7)$$

zusammengesetzt:

$$X_{i+1} = \gamma (A_\gamma - C X_i)^{-1} B \quad \text{für } i=0,1,\dots,N-1 \quad (6.8)$$

Anstatt nun das Grundgebiet tatsächlich entsprechend zu diskretisieren, dann die Steifigkeitsmatrix zu bestimmen und anschließend zu kondensieren, genügt es natürlich aus der Steifigkeitsmatrix für die erste Substruktur R_0 die Teilmatrizen A_γ, B, C zu gewinnen und damit K_N und X_N direkt aus der Rekursion zu gewinnen.

K_N ist dabei nichts anderes als die Elementsteifigkeitsmatrix zum Macro-Element $\bigcup_{i=0}^N R_i$ und kann ohne weiteres als Substrukturmatrix an das Restgebiet Ω^* gekoppelt werden.

Ein wichtiger Punkt, der hier noch kurz angesprochen werden soll, ist der Einbau der essentiellen Randbedingungen. In der theoretischen Formulierung dieser Arbeit sind sie implizit in der Wahl des Raumes V_h enthalten, d.h. die Basisfunktionen γ_i (vgl. (1.2)) haben diese Bedingung bereits zu erfüllen.

In der Praxis der Finiten Elemente werden sie bekanntlich durch Festsetzen der entsprechenden Freiheitsgrade erzwungen. Prinzipiell sind dabei 2 Methoden gebräuchlich (z.B. ZIENKIEWICZ (1977), SCHWARZ (1980)) :

- 1.) Man erfüllt die Festsetzung exakt, indem man in der Steifigkeitsmatrix die entsprechenden Zeilen und Spalten durch Nullen ersetzt, in die Diagonale eine Eins einträgt und auch die rechte Seite entsprechend korrigiert.

Dieses Verfahren ist allerdings sehr umständlich und wird daher meist nur in speziellen Anwendungen benutzt.

- 2.) Man erfüllt die Festsetzung nur näherungsweise, indem man die entsprechende Bedingungsgleichung mit einem sehr großen Fak-

tor auf die Diagonale und rechte Seite addiert. Bei entsprechender Wahl des Faktors ist dann der damit gemachte Fehler i. allg. weit unter dem Verfahrensfehler und kann vernachlässigt werden.

In beiden Fällen wird jedenfalls einfach die Steifigkeitsmatrix und rechte Seite geeignet modifiziert.

Wenn wir also davon ausgehen, daß die Modifikation im System der 0. Substruktur R_0 , also in K_0 , F_0 bereits eingebaut ist und die Randbedingung für alle Substrukturen R_i analog ist, dann gelten die Gleichungen (6.1) bis (6.8) ohne Einschränkung auch bei entsprechenden essentiellen Randbedingungen.

Die einzige Forderung an diese Randbedingungen im Bereich des Macroelements $\bigcup_{i=1}^N R_i$ ist, daß auch sie "mit der affinen Abbildung verträglich", d.h. an einander entsprechenden Knoten gleich, sein müssen.

Zum Schluß dieses Abschnittes noch ein einfacher Satz, der die Eigenschaften der Ausgangsmatrix S_0 zu R_0 auf alle K_i überträgt:

(6.1) SATZ:

Wenn S_0 positiv definit und symmetrisch ist, so sind es auch alle K_i $i=0, \dots, N$.

Beweis:

Die positive Definitheit ergibt sich in einfacher Weise durch vollst. Induktion aus Lemma (2.1).

Die Symmetrie folgt ebenfalls per Induktion sofort aus der Rekursion, wenn man beachtet, daß aus der Voraussetzung $C=B^T$ folgt und daß mit K_i auch immer K_i^{-1} symmetrisch ist.

$$K_{i+1} = \underbrace{A\gamma}_{\text{sym}} - \underbrace{B^T K_i^{-1} B}_{\text{sym}}$$

q.e.d.

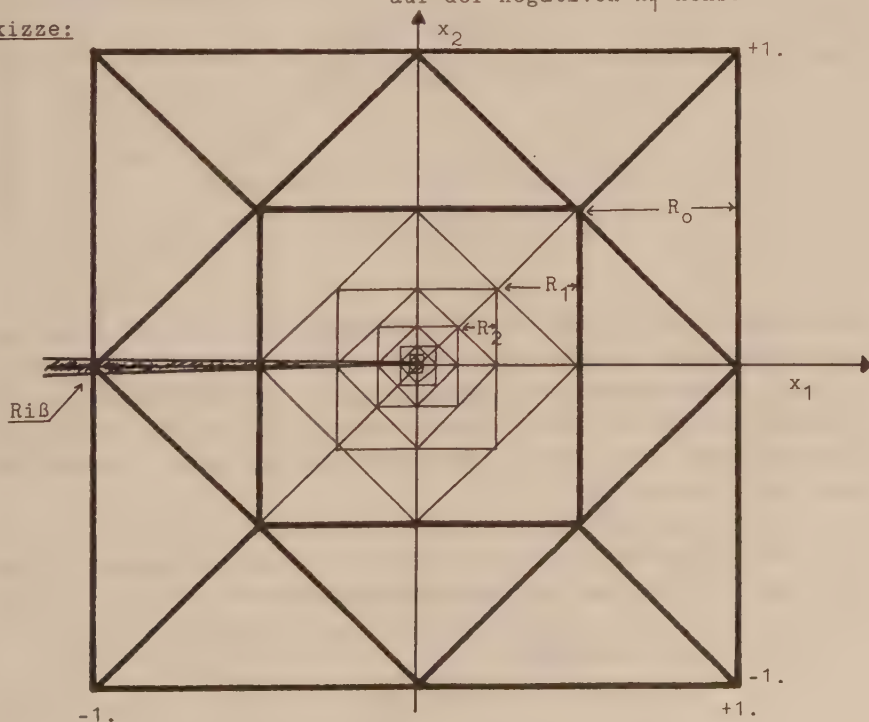
7) EIGENSCHAFTEN DER REKURSION

Wir wollen uns nun mit den Eigenschaften der Rekursionsformeln für $N \rightarrow \infty$ beschäftigen. Die Form des Macroelements $\bigcup_{i=0}^N R_i$ hängt dabei wesentlich von der Gestalt von R_0 und der affinen Transformation $\hat{T}$ ab. An 2 einfachen Beispielen soll das kurz demonstriert werden.

Wenn $\hat{T}$ beispielsweise eine zentrische Stauchung ist, so würde auf diese Weise eine unendliche Gitterverfeinerung um den Ursprung simuliert:

Beispiel: $\hat{T} \underline{x} = 1/2 \underline{x}$, $\Omega =$ Einheitsquadrat mit RiB auf der negativen x_1 -Achse

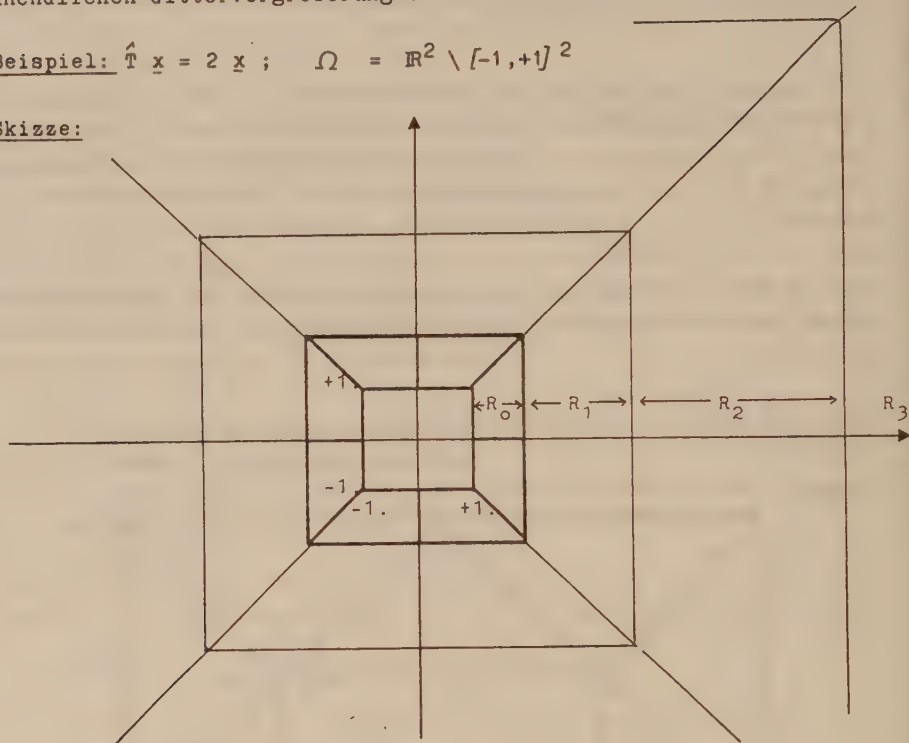
Skizze:



Wenn $\hat{T}$ zentrische Streckung ist, dann führt der Prozess zu einer unendlichen Gittervergrößerung :

Beispiel: $\hat{T} \underline{x} = 2 \underline{x}$; $\Omega = \mathbb{R}^2 \setminus [-1, +1]^2$

Skizze:



In diesem Abschnitt sei generell vorausgesetzt, daß die Gesamtmatrix symmetrisch aufgebaut ist, daß also $C=B^T$ gilt und A_γ sowie A_0 und A_u symmetrisch sind. Von A_γ , A_0 , A_u sei außerdem noch die Invertierbarkeit gefordert. (Wegen Lemma (2.1) ist das keine sehr einschneidende Bedingung). Die Matrix B muß weder symmetrisch noch invertierbar sein.

Die Rekursionsformel (6.2) kann man als Fixpunktiteration eines Operators $\mathcal{F}$ im Raum der $m \times m$ Matrizen ansehen:

$$\begin{aligned} \text{Sei } \mathcal{F} : D &\longrightarrow \mathbb{R}^{m \times m} \\ X &\longmapsto (A_\gamma - \gamma B X)^{-1} B^T \end{aligned}$$

Genauso definiert man für die Kondensation "von oben nach unten" einen Operator $\mathcal{G}(X) := \gamma(A_\gamma - B^T X)^{-1} B$.

Der Definitionsbereich $D \subset \mathbb{R}^{m \times m}$ von $\mathcal{F}$ kann dabei charakterisiert werden durch

$$D := \{X \in \mathbb{R}^{m \times m} \mid (A_\gamma - \gamma B X) \text{ ist invertierbar}\}$$

(und eine analoge Konstruktion für $\mathcal{G}$).

Zur weiteren Verwendung soll hier kurz an ein paar wichtige Begriffe aus der Matrixtheorie erinnert werden. Für weitergehende Einzelheiten sei z.B. auf das Buch von ZURMÜHL (1964) verwiesen.

(7.1) DEFINITIONEN:

Eine Matrixnorm $\|\cdot\|_M$ heißt mit einer Vektornorm $\|\cdot\|_V$ verträglich, wenn gilt:

$$\forall A \in \mathbb{R}^{m \times m}, x \in \mathbb{R}^m : \|Ax\|_V \leq \|A\|_M \|x\|_V$$

Eine Matrixnorm heißt submultiplikativ, wenn

$$\forall A, B \in \mathbb{R}^{m \times m} : \|AB\| \leq \|A\| \|B\|$$

Eine besondere Stellung in Bezug auf eine gegebene Vektornorm hat jeweils die sogenannte lub-Norm, definiert als:

$$\text{lub}(A) = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

Die lub-Norm ist stets verträglich und submultiplikativ.

Die lub-Norm zur EUKLIDischen Norm $\|x\|_2 = \sqrt{x^T x}$ ist die sogenannte HILBERT- oder Spektralnorm:

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$$

Der Spektralradius $\rho(A)$ einer Matrix A ist:

$$\rho(A) := \max_i |\lambda_i|$$

wobei $\lambda_i, i=1, \dots, m$ alle Eigenwerte von A sind.

Eine Matrix A heit normal :<===> $A^T A = A A^T$

In den folgenden Ausfhrungen gehen wir zunchst von einer beliebigen Vektornorm und einer damit vertrglichen, submultiplikativen Matrixnorm aus. Da eine Verwechslung kaum mglich ist, wollen wir sie beide formal durch $\|\cdot\|$ symbolisieren.

Nach diesen Vorbereitungen soll zunchst einmal die Menge D etwas handlicher gekennzeichnet werden. Das folgende Lemma gibt ein hinreichendes Kriterium fr $X \in D$.

(7.2) LEMMA:

$$\text{Es gilt: } \|X\| < \frac{1}{|\gamma| \|A_Y^{-1} B\|} \implies X \in D$$

Beweis:

Annahme: $X \notin D$

$$\langle == \rangle \quad A_Y - \gamma B X \quad \text{nicht invertierbar}$$

$$\langle == \rangle \quad \exists v \in \mathbb{R}^m, v \neq 0 : (A_Y - \gamma B X)v = 0$$

$$\langle == \rangle \quad A_Y v = \gamma B X v$$

$$\implies \quad \|v\| \leq |\gamma| \|A_Y^{-1} B\| \|X\| \|v\|$$

$$\langle == \rangle \quad \frac{1}{|\gamma| \|A_Y^{-1} B\|} \leq \|X\| \quad \text{im Widerspruch zur Voraussetzung}$$

q.e.d.

Von entscheidender Bedeutung für das ganze Verfahren ist natürlich die Frage nach der Konvergenz der Iterationen (6.2) bzw. (6.8). Also ob und wann gilt, daß

$$\|X_{i+1} - X_i\| \rightarrow 0 \quad \text{für } i \rightarrow \infty$$

Um dieses Problem zu klären, charakterisiert der folgende Satz eine Menge Γ , in der $\mathcal{F}$ kontrahierend ist, also mit $\kappa < 1$ gilt:

$$\text{für alle } X, Y \in \Gamma : \|\mathcal{F}(X) - \mathcal{F}(Y)\| \leq \kappa \|X - Y\|$$

Da alle Normen im $\mathbb{R}^{m \times m}$ äquivalent sind, ist es dabei zunächst egal auf welche Norm sich die Aussagen beziehen. Aus der Kontraktionseigenschaft folgt dann sofort (vgl. 7.4) die Konvergenz der Rekursion.

(7.3) SATZ:

$$\text{Sei } 0 < \gamma \leq 1; \|A_Y^{-1}B\| < \frac{1}{1+\gamma} \quad \text{und} \quad \|A_Y^{-1}B^T\| < \frac{1}{1+\gamma}$$

Dann bleibt die Rekursion $X_{i+1} = \mathcal{F}(X_i)$ in der Menge

$$\Gamma := \{X \in \mathbb{R}^{m \times m} / \|X\| \leq 1\}$$

und ist dort kontrahierend.

Beweis:

wir bemerken zunächst, daß aus $\|A_Y^{-1}B\| < \frac{1}{1+\gamma}$ sofort folgt

$$\frac{1}{\gamma \|A_Y^{-1}B\|} > \frac{1+\gamma}{\gamma} > 1$$

somit ist also jedenfalls $\Gamma \subset D$ wegen vorigem Lemma.

Teil 1: zeige: $X \in \Gamma \Rightarrow \mathcal{F}(X) \in \Gamma$

$$\begin{aligned}
\|\mathcal{F}(X)\| &= \|(A_Y - \gamma BX)^{-1} B^T\| = \\
&= \|(I - \gamma A_Y^{-1} BX)^{-1} A_Y^{-1} B^T\| \leq \\
&\leq \|(I - \underbrace{\gamma A_Y^{-1} BX}_{=: F})^{-1}\| \|A_Y^{-1} B^T\| < \quad (*)
\end{aligned}$$

Nebenrechnung:

Es ist $\|F\| = \|\gamma A_Y^{-1} BX\| \leq |\gamma| \|A_Y^{-1} B\| \|X\| < \frac{\gamma}{1+\gamma} < 1$
 also nach STÖR, BULIRSCH II (1973); Lemma auf Seite 153

$$\|(I - \gamma A_Y^{-1} BX)^{-1}\| \leq \frac{1}{1 + \frac{\gamma}{1+\gamma}} = 1 + \gamma$$

damit:

$$\dots (*) < (1+\gamma) \frac{1}{(1+\gamma)} = 1.$$

Teil 2: zeige: $X, Y \in \Gamma \implies \|\mathcal{F}(X) - \mathcal{F}(Y)\| < \kappa \|X - Y\|$
 mit $\kappa < 1$

Zur Abkürzung sei:

$$\kappa_1 := (1+\gamma) \|A_Y^{-1} B^T\| < 1$$

$$\kappa_2 := (1+\gamma) \|A_Y^{-1} B\| < 1$$

$$\begin{aligned}
\|\mathcal{F}(X) - \mathcal{F}(Y)\| &= \|(A_Y - \gamma BX)^{-1} B^T - (A_Y - \gamma BY)^{-1} B^T\| = \\
&= \|(A_Y - \gamma BX)^{-1} \{ (A_Y - \gamma BY) - (A_Y - \gamma BX) \} (A_Y - \gamma BY)^{-1} B^T\| = \\
&= \|(A_Y - \gamma BX)^{-1} B \gamma (X-Y) (A_Y - \gamma BY)^{-1} B^T\| \leq \\
&\leq \underbrace{\|(A_Y - \gamma BX)^{-1} B\|}_{\leq \kappa_2 \text{ analog Teil 1}} |\gamma| \|X-Y\| \underbrace{\|(A_Y - \gamma BY)^{-1} B^T\|}_{\leq \kappa_1 \text{ nach Teil 1}} = \\
&= \kappa \|X-Y\|
\end{aligned}$$

$$\text{mit } \kappa := \gamma \kappa_1 \kappa_2 < 1$$

q.e.d.

(7.4) KOROLLAR:

Die Voraussetzungen zu Satz (7.3) seien erfüllt. Ferner gelte

$$\|A_u^{-1}B\| \leq 1$$

Dann hat die Abbildung $\mathcal{T}$ genau einen Fixpunkt X_∞ in Γ und die Rekursion (6.3) konvergiert dagegen.

$$X_i \longrightarrow X_\infty \quad \text{für } i \longrightarrow \infty.$$

Beweis:

Nach Satz (7.3) ist $\mathcal{T} : \Gamma \longrightarrow \Gamma$ eine Kontraktion und Γ abgeschlossen. Nach Voraussetzung liegt der Startwert $A_u^{-1}B$ der Rekursion in Γ .

Aus dem Kontraktionssatz von BANACH (siehe z.B. WLOKA (1971)) folgt somit sofort die Behauptung.

q.e.d.

In völliger Analogie lassen sich ähnliche Sätze auch für die Kondensation "von oben nach unten" (Operator $\mathcal{G}$) und die damit verbundene Rekursion (6.8) beweisen. Es ändert sich dabei nur die Forderung an den Verträglichkeitsfaktor γ . Da die Argumentation gleich bleibt, soll das hier aber nicht explizit ausgeführt, sondern nur das Ergebnis angegeben werden:

(7.5) SATZ:

$$\text{Sei } \gamma \geq 1; \|A_Y^{-1}B\| < \frac{1}{1+\gamma} \quad \text{und} \quad \|A_Y^{-1}B^T\| < \frac{1}{1+\gamma}$$

Dann bleibt die Rekursion $X_{i+1} = \mathcal{G}(X_i)$ in der Menge

$$\Gamma := \{X \in \mathbb{R}^{m \times m} \mid \|X\| \leq 1\}$$

und ist dort kontrahierend.

Beweis: analog zum Beweis von Satz (7.3) .

Da der Fixpunkt eindeutig ist in Γ , ist die Startmatrix für die Konvergenz der Iteration völlig unerheblich. Das bedeutet, daß als wesentliche Voraussetzung für Konvergenz übrigbleibt:

$$\|A_{\gamma}^{-1}B\| < \frac{1}{1+\gamma} \quad \text{und} \quad \|A_{\gamma}^{-1}B^T\| < \frac{1}{1+\gamma}$$

in irgendeiner submultiplikativen Norm. Diese Bedingung muß im Einzelfall überprüft werden. In numerischen Experimenten war sie bei allen untersuchten Beispielen tatsächlich erfüllt, wenn als Norm die HILBERT-Norm zugrunde gelegt wurde.

Ein allgemeiner Nachweis konnte nicht gefunden werden. Aber der folgende Satz zeigt immerhin, daß die Spektralradien der beiden fraglichen Matrizen nicht allzu groß werden können.

(7.6) SATZ:

Sei $\gamma > 0$. Dann gilt:

$$\rho(A_{\gamma}^{-1}B) = \rho(A_{\gamma}^{-1}B^T) \leq \frac{1}{1+\gamma}$$

(Das Gleichheitszeichen kann dabei höchstens dann eintreten, wenn in A_{γ} und B keine essentiellen Randbedingungen eingebaut sind (vgl. Seite 57)).

Beweis:

1. zeige: $A_{\gamma}^{-1}B$ und $A_{\gamma}^{-1}B^T$ haben die gleichen Eigenwerte
Das charakteristische Polynom von $A_{\gamma}^{-1}B$ ist gegeben durch:

$$\begin{aligned} p_{A_{\gamma}^{-1}B}(\lambda) &= \det(A_{\gamma}^{-1}B - \lambda I) = \\ &= \det(B - \lambda A_{\gamma}) \det(A_{\gamma}^{-1}) = \quad (\text{weil } A_{\gamma} \text{ sym.}) \\ &= \det(B^T - \lambda A_{\gamma}) \det(A_{\gamma}^{-1}) = \\ &= \det(A_{\gamma}^{-1}B^T - \lambda I) = \\ &= p_{A_{\gamma}^{-1}B^T}(\lambda) \end{aligned}$$

dies ist gerade das charakteristische Polynom von $A_{\gamma}^{-1}B^T$.
 Ergo haben $A_{\gamma}^{-1}B$ und $A_{\gamma}^{-1}B^T$ die gleichen Eigenwerte und
 folglich auch den gleichen Spektralradius.

2. zeige: λ Eigenwert von $A_{\gamma}^{-1}B \implies 1/\lambda \leq \frac{1}{1+\gamma}$

Sei $0 \neq x \in \mathbb{R}^m$ beliebig vorgegeben. Und sei:

$$Z := (\underbrace{x, x, \dots, x, x^*}_{N \text{ Blöcke}}) \in \mathbb{R}^{Nm}$$

Dabei soll x^* in den ersten Komponenten mit x übereinstimmen, aber im unteren Teil nur Nullen enthalten - gerade so, daß die in (5.6) eingeführten Dummyvariablen wieder ausgeblendet werden.

Die Gesamtsteifigkeitsmatrix K_{ges} einer N-fachen Verschachtelung von Substrukturen gemäß (5.3) kann nach Lemma (2.1) jedenfalls als positiv semidefinit angenommen werden. Wenn die essentiellen Randbedingungen eingebaut sind sogar als positiv definit. Folglich gilt auch für die umgeformte Matrix nach (5.6):

$$0 \leq Z^T \begin{pmatrix} A_0 & \gamma B \\ B^T & A & \gamma B \\ & \ddots & \ddots & \ddots \\ & & B^T & \gamma B \\ & & & A_u \end{pmatrix} Z =$$

$$\begin{aligned} &= x^T A_0 x + x^T B x + \\ &+ x^T B^T x + x^T A_{\gamma} x + \gamma x^T B x + \\ &+ \dots \\ &+ \dots \\ &+ x^T B^T x + x^T A_{\gamma} x + \gamma x^T B x^* + \\ &+ x^{*T} B^T x + x^{*T} A_u x^* \end{aligned}$$

wegen der speziellen Form der Matrizen (siehe Seite 51/52) kann man das zusammenfassen zu:

$$\begin{aligned} &= (N-1) (x^T B^T x + x^T A_\gamma x + \gamma x^T B x) = \\ &= (N-1) (x^T A_\gamma x + (1+\gamma) x^T B x) . \end{aligned}$$

Daraus folgt sofort:

$$\frac{x^T B x}{x^T A_\gamma x} \geq \frac{1}{1+\gamma}$$

Völlig analog kann mit $Z := (x, -x, +x, -x, \dots)$ gefolgert werden, daß

$$\frac{x^T B x}{x^T A_\gamma x} \leq \frac{1}{1+\gamma}$$

gelten muß. Also gilt insgesamt für den verallgemeinerten RAYLEIGH-Quotienten die Abschätzung:

$$\left| \frac{x^T B x}{x^T A_\gamma x} \right| \leq \frac{1}{1+\gamma}$$

Sei nun x ein Eigenvektor von $A_\gamma^{-1} B$ zum Eigenwert λ dann folgt sofort:

$$|\lambda| = \left| \frac{x^T B x}{x^T A_\gamma x} \right| \leq \frac{1}{1+\gamma}$$

Damit gilt also auch für den Spektralradius

$$\rho(A_\gamma^{-1} B) = \max_{i=1, \dots, m} |\lambda_i| \leq \frac{1}{1+\gamma}$$

q.e.d.

(7.7) KOROLLAR:

Wenn $A_Y^{-1}B$ und $A_Y^{-1}B^T$ normal sind und die essentiellen Randbedingungen enthalten, dann gilt

$$\|A_Y^{-1}B\|_2 = \|A_Y^{-1}B^T\|_2 < \frac{1}{1 + \gamma}$$

wobei $\|\cdot\|_2$ die HILBERT-Norm ist.

Beweis:

für normale Matrizen M gilt $\|M\|_2 = \rho(M)$ womit aus Satz (7.6) sofort die Behauptung folgt.

q.e.d.

Die Rekursionsformeln für die rechte Seite H_i (siehe (6.2) bzw. (6.7)) lassen sich nicht in so einfacher Weise als Fixpunktiteration interpretieren, da sich die Matrix K_i ja von Schritt zu Schritt ändert. Dafür läßt sich hier aber der Grenzwert H_∞ - wenn er existiert - sehr leicht angeben:

(7.8) LEMMA:

Die Iteration (6.2) bzw. (6.8) habe einen Fixpunkt mit H_∞ und K_∞ . Dann gilt

$$H_\infty = (I + BK_\infty^{-1})^{-1} F \quad \text{bei (6.2)}$$

$$H_\infty = (I + CK_\infty^{-1})^{-1} F \quad \text{bei (6.7)}$$

Beweis:

Für einen Fixpunkt von (6.2) gilt:

$$H_\infty = F - BK_\infty^{-1} H_\infty \quad \text{woraus sofort die Beh. folgt.}$$

analog für (6.7). *

KAPITEL 3: DIE FIXPUNKTGLEICHUNG

8) MATRIXPOLYNOME UND LATENTE WURZELN

Für den Fixpunkt X_∞ , gegen den die Rekursion des vorigen Abschnittes konvergiert, gilt:

$$\mathcal{F}(X_\infty) = X_\infty \quad (8.1)$$

also ausgeschrieben und nach Multiplikation mit $-(A_Y - \gamma B X_\infty)$ von links:

$$\gamma B X_\infty^2 - A_Y X_\infty + C = 0 \quad (8.2)$$

Für den Fixpunkt Y_∞ , der aus der Rekursion mit dem Operator $\mathcal{G}$ stammt, gilt analog, aber mit vertauschten Koeffizienten:

$$C Y_\infty^2 - A_Y Y_\infty + \gamma B = 0 \quad (8.3)$$

Die Fixpunkte ergeben sich also gerade als Nullstellen von bestimmten quadratischen Matrixpolynomen. Zwischen den Nullstellen von (8.2) und denen von (8.3) besteht dabei eine enge Verwandtschaft:

(8.1) LEMMA:

Sei $X \in \mathbb{R}^{m \times m}$ eine invertierbare Lösung von (8.2).

Dann ist $Y := X^{-1}$ eine Lösung von (8.3).

Beweis:

aus $\gamma B X^2 - A_Y X + C = 0$ folgt durch Multiplikation von rechts mit X^{-2} : $\gamma B - A_Y X^{-1} + C X^{-2}$, also gerade die Behauptung.

q.e.d.

Mit dem Ziel die gesuchte Matrix X_∞ bzw. Y_∞ eventuell direkt aus Gleichung (8.2) bzw. (8.3) bestimmen zu können, soll nun die Struktur der Nullstellen von quadratischen Matrixpolynomen. näher untersucht werden.

Derartige Polynome treten zum Beispiel auch bei der Analyse von gedämpften Schwingungen auf. Die Monographie von LANCASTER (1966) kann hier wohl als Standardwerk zitiert werden. Auch in den Büchern von MacDUFFEE (1946), GANTMACHER (1960) und ZURMÜHL (1964) finden sich bereits viele grundlegende Aussagen.

Eine umfassende Theorie für allgemeine Matrixpolynome beliebiger Ordnung wurde aber erst 1976 von DENNIS, TRAUB und WEBER veröffentlicht. Sie beschränken sich dabei allerdings auf sogenannte monische (engl.: monic) Polynome, d.h. Polynome, deren Koeffizient höchster Ordnung die Einheitsmatrix ist. Solange dieser Koeffizient invertierbar ist, ist dies ja keine schwerwiegende Einschränkung.

Wir wollen die wichtigsten Ideen am Beispiel eines allgemeinen Matrixpolynoms $P(X)$ 2. Ordnung hier kurz zusammenstellen.

Sei also:

$$\begin{array}{ccc} P : \mathbb{R}^{m \times m} & \longrightarrow & \mathbb{R}^{m \times m} \\ X & \longmapsto & P(X) = RX^2 + SX + T \end{array}$$

$$\text{mit: } R, S, T \in \mathbb{R}^{m \times m}$$

In engem Zusammenhang mit dem Matrixpolynom $P(X)$ steht die sogenannte Lambda-Matrix zu $P(X)$. Sie ist definiert als:

$$P(\lambda I) = \lambda^2 R + \lambda S + T \quad (8.4)$$

($I = m \times m$ -Einheitsmatrix)

Weitere wichtige Begriffe sind latente Wurzeln (engl.: latent roots) und latente Vektoren (engl.: latent vector), welche in gewisser Weise das Konzept von Eigenwert-Eigenvektor verallgemeinern:

(8.2) DEFINITION:

λ heißt latente Wurzel des Matrixpolynoms $P(X)$, wenn gilt:

$P(\lambda I)$ ist singulär.

Ein Vektor $b \in \mathbb{R}^m$ $b \neq 0$ mit $P(\lambda I) \cdot b = 0$ bzw. $b^T P(\lambda I) = 0$ heißt dann rechter (bzw. linker) latenter Vektor dazu.

Analog zu den Begriffen bei Eigenwerten ist die algebraische Vielfachheit von λ die Vielfachheit als Nullstelle von $\det(P(\lambda I))$ und die geometrische Vielfachheit die Dimension des Kerns von $P(\lambda I)$.

Wie man sich leicht klarmacht (vollständige Induktion nach m !) gilt für die Determinante der Lambda-Matrix die Entwicklung:

$$\det(P(\lambda I)) = \lambda^{2m} \det(R) + \dots + \det(T)$$

Somit kann es maximal $2m$ latente Wurzeln zu $P(X)$ geben. Wenn R nichtsingulär ist, dann gibt es genau $2m$ latente Wurzeln (mit Vielfachheit gezählt). In diesem Fall läßt sich $P(X)$ durch Multiplikation mit R^{-1} von links auf monische Form bringen.

Da R, S, T als reell vorausgesetzt waren, also $\det(P(I))$ ein reelles Polynom ist, folgt auch sofort, daß die latenten Wurzeln entweder rein reell sind oder in Gestalt von konjugiert komplexen Paaren $\lambda, \bar{\lambda}$ auftreten.

Diese $2m$ latenten Wurzeln mit zugehörigen latenten Vektoren lassen sich beispielsweise aus dem folgenden verallgemeinerten Ei-

genwertproblem der Dimension $2m \times 2m$ berechnen:

$$\begin{pmatrix} S & T \\ I & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} -R & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (8.5)$$

Aus der 2. Blockzeile sieht man sofort, daß

$$x = \lambda y \quad (8.6)$$

gelten muß und dies ergibt eingesetzt in die 1. Blockzeile

$$\lambda^2 R y + \lambda S y + T y = 0 \quad (8.7)$$

also gerade die Bedingung dafür, daß λ latente Wurzel bzw. y latenter Vektor von $P(X)$ sind.

Für die effiziente Lösung solcher verallgemeinerter Eigenwertprobleme stehen bekannte Algorithmen zur Verfügung, etwa das QZ-Verfahren von MOLER und STEWARD (1973). Insbesondere ist dafür die Invertierbarkeit der beiden großen Blockmatrizen in (8.5) nicht erforderlich.

Den Zusammenhang zwischen latenten Wurzeln, latenten Vektoren und eigentlich gesuchten Nullstellen X des Polynoms $P(X)$ gibt nun das folgende einfache Lemma:

(8.3) LEMMA:

Sei $X \in \mathbb{R}^{m \times m}$ und $P(X) = 0$.

Seien λ, v Eigenwert bzw. rechter/linker Eigenvektor von X . Dann sind λ, v auch latente Wurzel bzw. rechter/linker latenter Vektor der Lambda-Matrix zu P .

Beweis: Sei $Xv = \lambda v$

$$\begin{aligned} \Rightarrow 0 &= 0 \cdot v = P(X) \cdot v = \\ &= RX^2v + SXv + Tv = \\ &= \lambda^2 Rv + \lambda Sv + Tv = \\ &= P(\lambda I) v \end{aligned}$$

Analog für linke Eigenvektoren.

q.e.d.

In gewisser Weise gilt auch die Umkehrung des letzten Lemmas. Daraus ergibt sich ein konstruktives Verfahren, um aus den latenten Wurzeln und Vektoren, also den entsprechenden Komponenten des Eigensystems von (8.5), eine Nullstelle von $P(X)$ zu erzeugen:

(8.4) SATZ:

Jede Auswahl von m linear unabhängigen latenten Vektoren v_i $i=1, \dots, m$ mit zugehörigen latenten Wurzeln λ_i liefert eine Nullstelle von $P(X)$ durch:

$$X = V D V^{-1}$$

$$\begin{aligned} \text{mit } V &:= (v_1, v_2, \dots, v_m) \\ D &:= \text{diag}(\lambda_i) \end{aligned}$$

BEWEIS:

Nach Lemma (8.3) sind die Eigenwerte/Eigenvektoren auch latente Wurzeln/ Vektoren von P , also:

$$\lambda_i^2 Rv_i + \lambda_i Sv_i + Tv_i = 0 \quad i=1, \dots, m$$

Schreibt man alle diese Spalten nebeneinander, so folgt:

$$RV \text{diag}(\lambda_i^2) + SV \text{diag}(\lambda_i) + TV = 0$$

da die v_i linear unabhängig sind, ist V invertierbar und es folgt nach Multiplikation von rechts mit V^{-1} :

$$RV\text{diag}(\lambda_i^2)V^{-1} + SV\text{diag}(\lambda_i)V^{-1} + TVV^{-1} = 0$$

$$\text{also } RX^2 + SX + T = 0$$

q.e.d.

Offenbar muß also zumindest jede diagonalisierbare Polynomnullstelle X von der Form VDV^{-1} sein, wobei die Spalten von V gerade rechte latente Vektoren zu $P(X)$ sind und D die latenten Wurzeln in der Diagonale enthält. Da es maximal $2m$ latente Wurzeln gibt, kann es also höchstens $\binom{2m}{m}$ Möglichkeiten geben, davon m Stück auszuwählen. Für Polynome mit invertierbarem höchsten Koeffizienten läßt sich diese Aussage zu einem Kriterium über die Gesamtzahl der Nullstellen präzisieren:

(8.5) KOROLLAR

Sei R invertierbar.

Wenn $P(X)$ $2m$ verschiedene latente Wurzeln hat und die zugehörigen rechten latenten Vektoren die HAAR-Bedingung)* erfüllen, dann hat $P(X)$ genau $\binom{2m}{m}$ verschiedene Nullstellen.

Beweis:

Multipliziert man $P(X)$ von links mit R^{-1} , so ist das Polynom monisch und es folgt mit DENNIS, TRAUB, WEBER (1976), Corollary 4.1, Seite 837 sofort die Behauptung.

q.e.d.

Ein Matrixpolynom hat also in aller Regel erheblich mehr Nullstellen, als ein skalares Polynom gleichen Grades.

)* Ein Satz von Vektoren v_i $i=1,2,\dots$ erfüllt die HAAR-Bedingung : $\Leftrightarrow$ Jede Auswahl von m Vektoren ist linear unabhängig.

Bei den speziellen Matrixpolynomen definiert in Gleichung (8.2) bzw. (8.3) muß davon ausgegangen werden, daß die Matrizen B und C (siehe Seite 51/52) auch singulär werden können. In diesem Fall ist die Theorie von DENNIS, TRAUB und WEBER nicht mehr anwendbar. Unter den latenten Werten tauchen dann unendliche Größen auf.

Die naive Betrachtung infinitesimaler latenter Werte ist natürlich etwas fragwürdig. Man muß das so verstehen, daß bei einer kleinen Störung der Matrix R, die die Singularität aufhebt, endliche Eigenwerte entstehen, die bei stetiger Reduzierung der Störung unbeschränkt wachsen.

In der Praxis sieht das so aus, daß z.B. der QZ-Algorithmus von MOLER und STEWARD zu jedem Eigenwert λ_i von (8.5) ein Wertepaar (a_i, b_i) ausgibt. Die Zuordnung $\lambda_i = a_i/b_i$ liegt dann beim Anwender. Für $b_i = 0$ wird λ_i formal als unendlich betrachtet.

(8.6) LEMMA:

Sei $P(X) = RX^2 + SX + T$. Dann gilt:

- i) $\text{rang}(T) = m - m_1 \iff \lambda = 0$ ist latente Wurzel mit geometrischer Vielfachheit m_1
- ii) $\text{rang}(R) = m - m_2 \iff \lambda = \infty$ ist latente Wurzel mit geometrischer Vielfachheit m_2

Beweis:

für $\lambda = 0$ ist $P(\lambda I) = T$, hat also nach Voraussetzung einen m_1 -dimensionalen Nullraum.

Mit $\tau := \frac{1}{\lambda}$ ist:

$P(\lambda I) = \frac{1}{\tau^2}R + \frac{1}{\tau}S + T$ singulär $\iff R + \tau S + \tau^2 T$ singulär
Für $\tau \rightarrow 0$ entspricht dies dem Fall i).

q.e.d.

(8.7) LEMMA:

Sei $P(X) = RX^2 + SX + T$ mit $R = T^T$ und $S = S^T$.

Dann gilt:

λ latente Wurzel von $P \iff 1/\lambda$ latente Wurzel von P

Beweis:

λ latente Wurzel von $P \iff (\lambda^2 R + \lambda S + T)$ singular
 $\iff$ es gibt linken latenten Vektor $w \in \mathbb{R}^m$ mit:

$$w^T (\lambda^2 R + \lambda S + T) = 0$$

$$\iff \lambda^2 w^T R + \lambda w^T S + w^T T = 0$$

transponiert man diese Gleichung, so folgt wegen $S=S^T$

$$\lambda^2 R^T w + \lambda S w + T^T w = 0$$

nach Division durch λ^2 :

$$R^T w + \frac{1}{\lambda} S w + \frac{1}{\lambda^2} T^T w = 0$$

wenn man nun $R=T^T$ berücksichtigt, so ist dies aber gerade wieder $P(1/\lambda \cdot I)w=0$, woraus die Behauptung folgt.

q.e.d.

Dieses Lemma gilt formal auch für Paare $0, \infty$ von latenten Werten.

Solange keine essentiellen Randbedingungen eingebaut sind (vgl. Seite 57/58), sind die Steifigkeitsmatrizen aus der Finite-Elemente-Methode bekanntlich noch singular. Dies entspricht normalerweise dem konstanten Potential bei Potentialproblemen, bzw. den Starrkörperverschiebungen in der Elastizitätstheorie. Daraus ergibt sich folgende Aussage:

(8.8) LEMMA:

Die Gesamtsteifigkeitsmatrix zu einer N-fachen Verschachtelung von Substrukturen gemäß (5.3) habe den Eigenwert 0 zu 1 verschiedenen Eigenvektoren $e_i = (v_i, v_i, \dots, v_i)$ $i=1, \dots, 1$

Dann gilt:

$\lambda = -1$ ist (geometrisch) 1-fache latente Wurzel zu den Matrixpolynomen aus (8.2) bzw. (8.3).

Beweis:

Schreibt man die Voraussetzung mit der Matrix aus Gleichung (5.4) aus und multipliziert die Blockzeilen jeweils mit $\gamma^{2 \cdot i}$ wie auf Seite 51 beschrieben, dann steht gerade N mal die Gleichung

$$Cv_i + A_\gamma v_i + \gamma Bv_i = 0$$

da, woraus sich unmittelbar die Behauptungen ergeben.

q.e.d.

Wir wollen nun die volle Allgemeinheit verlassen. Ein sehr allgemeiner Fall liegt vor, wenn wir nur davon ausgehen, daß die zugrundegelegte Bilinearform $a(.,.)$ (siehe Seite 8) symmetrisch und unter der verwendeten affinen Transformation invariant (vgl. §4) ist.

In diesem Fall, der in der Praxis sehr häufig ist, haben wir:

$$\gamma = 1, \quad C = B^T, \quad A_\gamma = A_\gamma^T.$$

Der Rangdefekt von B sei k (also $\text{rang}(B) = m - k$ mit $0 \leq k < m$) und in der Gesamtsteifigkeitsmatrix mögen noch l essentielle Randbedingungen fehlen (vgl. Lemma(8.8)). Nimmt man dann die Aussagen der Lemmata (8.6), (8.7) und (8.8) zusammen, so ergibt sich folgende Struktur wenn man die latenten Wurzeln ihrer absoluten Größe nach

anordnet:

m Stück dem Betrag nach ≤ 1		m Stück dem Betrag nach ≥ 1	
$\lambda_1 = -1$ $\vdots$ $\lambda_l = -1$	1 Stück nach Lemma (8.8)	$\lambda_{m+1} = -1$ $\vdots$ $\lambda_{m+l} = -1$	
λ_{l+1} $\vdots$ λ_{m-k}	nach Lemma (8.7)	$\lambda_{m+l+1} = 1/\lambda_{l+1}$ $\vdots$ $\lambda_{2m-k} = 1/\lambda_{m-k}$	
$\lambda_{m-k+1} = 0$ $\vdots$ $\lambda_m = 0$	k Stück nach Lemma (8.6)	$\lambda_{2m-k+1} = \infty$ $\vdots$ $\lambda_{2m} = \infty$	

(8.9) SATZ:

Es gelte obige Anordnung der latenten Werte. Die zur linken Hälfte gehörigen latenten Vektoren v_i $i=1, \dots, m$ seien linear unabhängig.

Die Fixpunkte X_∞ bzw. Y_∞ von $\mathcal{F}$ bzw. $\mathcal{G}$ seien diagonalähnlich.

Dann gilt:

$$X_\infty = Y_\infty = VDV^{-1}$$

$$\text{mit: } V = (v_1, v_2, \dots, v_m)$$

$$D = \text{diag}(\lambda_i) \quad i=1, \dots, m$$

Beweis:

Wenn X_∞ diagonalähnlich ist, dann gibt es $\hat{D}, \hat{V}$ mit

$$X_\infty = \hat{V}\hat{D}\hat{V}^{-1}$$

wobei $\hat{D}$ die Diagonalmatrix der Eigenwerte und $\hat{V}$ die Spaltenmatrix der Eigenvektoren dazu ist.

Nach Lemma (8.3) muß jeder Eigenvektor/ Eigenwert von X_∞ latenter Vektor/ latente Wurzel des Matrixpolynoms sein, also in obiger Liste vorkommen.

Es gilt $\|X_\infty\| < 1$ wegen $X \in \Gamma$ (vgl. Satz (7.3)) und da für alle Eigenwerte λ und beliebige Normen stets $\|X_\infty\| \neq \lambda$ gilt, kommen als Eigenwerte von X_∞ nur noch die Werte der linken Spalte in Frage. Da diese genau m Werte enthält und die zugehörigen Vektoren linear unabhängig sind, muß (bis auf eine Permutation der Spalten) $V = \hat{V}$ und $D = \hat{D}$ gelten.

Genau die gleiche Argumentation gilt auch für Y_∞ .

q.e.d.

9) LÖSUNGSVERFAHREN

Wir wollen nun ein paar Verfahren zur numerischen Lösung der Matrixgleichungen des letzten Abschnittes diskutieren.

Das naheliegendste Verfahren zur Bestimmung der gesuchten Nullstelle von

$$\begin{aligned} \gamma BX^2 - A_\gamma X + C &= 0 \\ \text{bzw. } CY^2 - A_\gamma Y + \gamma B &= 0 \end{aligned} \quad (9.1)$$

ist natürlich die Originalrekursion

$$\begin{aligned} X_{i+1} &= (A_\gamma - \gamma BX_i)^{-1} C \\ \text{bzw. } Y_{i+1} &= \gamma (A_\gamma - CX_i)^{-1} B \end{aligned} \quad (9.2)$$

aus der die Gleichung jeweils abgeleitet wurde (vgl. §6). Vorteil bei dieser Methode ist, daß sie sicherlich den eigentlich gesuchten Fixpunkt, also die X-Matrix für eine infinite Substruktur-schachtelung ergibt.

Ein entscheidender Nachteil ist der große numerische Aufwand, den die Inversion der Matrix $(A_\gamma - \gamma BX_i)$ bzw. $(A_\gamma - CX_i)$ bei jedem Schritt verursacht.

HOUDE HAN (1982) geht im Prinzip diesen Weg. Er verwendet dabei direkt die 2-Schritt-Iteration, wie sie in (6.2) formuliert ist, ohne expliziten Konvergenzbeweis. Er berichtet, daß ca. 7-8 Schritte ausreichend sind für akzeptable Genauigkeit.

Nach unseren Erfahrungen mit den Formeln (9.2) ist die Konvergenz i. allg. linear und die Geschwindigkeit hängt sehr stark davon ab, wie klein der Kontraktionsmodul χ in Satz (7.3) bzw. (7.4) im speziellen Fall gerade ist.

Eine zweite Möglichkeit besteht darin, das verallgemeinerte Eigenwertproblem

$$\begin{pmatrix} A & C \\ I & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} \gamma B & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (9.3)$$

(bzw. mit C und γB vertauscht) zu lösen (vgl. Gleichung (8.5)) und aus den Komponenten dieser Lösungen gemäß Satz (8.9) dann den gesuchten Fixpunkt zu synthetisieren. Da man dazu aber praktisch das komplette Eigensystem eines doppelt so großen verallgemeinerten Eigenwertproblems benötigt, dürfte dieser Weg zumal bei großen Matrizen eher noch ungünstiger sein, als die Originalrekursion.

Nur im Sonderfall $\gamma=1$ und $B=C$ kann er sehr effektiv sein. Dann entartet das quadratische Eigenwertproblem nämlich in ein lineares. Dieser Fall ist möglich, tritt aber nur bei ganz speziellen F.E.-Zerlegungen in 2 Dimensionen auf.

In den Arbeiten von THATCHER (1975) (1976) wird nur dieser Spezialfall mit zusätzlich als invertierbar angenommenem B näher untersucht. In diesem Fall sind die beiden quadratischen Gleichungen (8.2) und (8.3) identisch und die latenten Wurzeln lassen sich recht einfach bestimmen:

Aus $\lambda^2 Bv - \lambda A_\gamma v + Bv = 0$ folgt sofort das verallgemeinerte lineare Eigenwertproblem

$$Bv = \tau A_\gamma v \quad (9.4)$$

$$\text{mit: } \tau := \frac{\lambda}{1+\lambda}$$

das mit bekannten Methoden gelöst werden kann. So ist es bei invertierbarem B äquivalent zum gewöhnlichen Eigenwertproblem

$$B^{-1} A_\gamma v = 1/\tau \cdot v \quad (9.5)$$

Aus jedem Eigenwert τ_i $i=1, \dots, m$ mit Eigenvektor v_i lassen sich dann je zwei latente Wurzeln $\lambda_i^{(1)}, \lambda_i^{(2)}$ zum latenten Vektor v_i bestimmen, gemäß:

$$\lambda_i^{1/2} = \frac{1}{2\tau_i} \pm \sqrt{\left(\frac{1}{2\tau_i}\right)^2 - 1} \quad (9.6)$$

Man beachte, daß auch hier der von Lemma (8.7) vorhergesagte Zusammenhang

$$\lambda_i^{(1)} = 1 / \lambda_i^{(2)}$$

gilt.

Die τ_i sind dann die Eigenwerte von $A_Y^{-1}B$. Da in obigem Spezialfall diese Matrix normal ist, ist die Forderung, die THATCHER an die τ_i stellt, gerade ein Spezialfall unserer Forderung $\|A_Y^{-1}B\| < \frac{1}{1+\gamma}$ aus Satz (7.3).

DENNIS, TRAUB und WEBER (1978) beschreiben im 2. Teil ihrer Arbeit verschiedene Algorithmen für die Nullstellenbestimmung bei monischen Polynomen beliebiger Ordnung. Diese sind Verallgemeinerungen entsprechender Algorithmen für skalare Polynome. Leider konvergieren sie gegen eine sogenannte dominante Lösung, d.h. gegen diejenige Polynomnullstelle mit den größten Eigenwerten - also genau die verkehrte für unsere Zwecke.

DAVIS (1983) hat eine FORTRAN-Routine für Lösungen der allgemeinen quadratischen Matrixgleichung $AX^2+BX+C=0$ veröffentlicht, die im Prinzip auf dem NEWTON-Verfahren beruht. Bei geeigneter Startnäherung sollte dieser Algorithmus quadratisch gegen den gewünschten Fixpunkt konvergieren. Der Programmieraufwand dafür ist allerdings nicht unerheblich.

In diesem Zusammenhang ist vielleicht auch die Arbeit von AFOLABI (1982) erwähnenswert, der durch direkte Formelmanipulation in der höheren Programmiersprache ALTRAN Matrizen mit polynomialen Koeff-

fizienten invertiert. Für die praktische Anwendung dürfte dieser Weg aber noch nicht reif sein.

Wir wollen hier nun ein weiteres Iterationsverfahren vorschlagen, das eine gewisse Ähnlichkeit mit dem Block-BERNOULLI-Algorithmus von DENNIS, TRAUB und WEBER hat, aber für unseren Spezialfall geeignet ist.

(9.1) SATZ

Sei $0 < \gamma \leq 1$, $\|A_\gamma^{-1}C\| < \frac{1}{1+\gamma}$ und $\|A_\gamma^{-1}B\| < \frac{1}{1+\gamma}$

Dann konvergiert die Rekursion

$$X_0 := 0$$

$$X_{i+1} := A_\gamma^{-1}C + \gamma A_\gamma^{-1}B X_i^2 \quad \text{für } i=1,2,3,\dots$$

gegen diejenige Lösung von $\gamma B X^2 - A_\gamma X + C = 0$ mit $\|X\| \leq 1$.

Beweis:

Die Rekursionsvorschrift ist die Fixpunktiteration zum Operator:

$$\begin{aligned} \mathcal{H} : \mathbb{R}^{m \times m} &\longrightarrow \mathbb{R}^{m \times m} \\ X &\longmapsto \mathcal{H}(X) := A_\gamma^{-1}C + \gamma A_\gamma^{-1}B X^2 \end{aligned}$$

Genau wie bei Satz (7.3) setzen wir

$$\Gamma := \{X \mid \|X\| \leq 1\}$$

a.) Es gilt: $\mathcal{H} : \Gamma \longrightarrow \Gamma$

$$\begin{aligned} \|\mathcal{H}(X)\| &= \|A_\gamma^{-1}C + \gamma A_\gamma^{-1}B X^2\| \leq \\ &\leq \|A_\gamma^{-1}C\| + \gamma \|A_\gamma^{-1}B\| \cdot \|X\|^2 < \end{aligned}$$

$$\leq \frac{1}{1+\gamma} + \frac{\gamma}{1+\gamma} = 1$$

b.) für $X, Y \in \Gamma$ ist

$$\begin{aligned} \| \mathcal{K}(X) - \mathcal{K}(Y) \| &= \| \gamma A_{\gamma}^{-1} B (X^2 - Y^2) \| \leq \\ &\leq \gamma \| A_{\gamma}^{-1} B \| \| X^2 - Y^2 \| = \\ &= \gamma \| A_{\gamma}^{-1} B \| \| 1/2 (X+Y)(X-Y) + 1/2 (X-Y)(X+Y) \| \leq \\ &\leq \gamma \| A_{\gamma}^{-1} B \| (1/2 \| X+Y \| \| X-Y \| + 1/2 \| X-Y \| \| X+Y \|) = \\ &= \gamma \| A_{\gamma}^{-1} B \| \| X+Y \| \| X-Y \| \leq \\ &\leq \gamma \| A_{\gamma}^{-1} B \| 2 \| X-Y \| \end{aligned}$$

Setzen wir also $\alpha := 2 \cdot \| A_{\gamma}^{-1} B \| < \frac{2\gamma}{1+\gamma} \leq 1$

so folgt, daß $\mathcal{K}$ in Γ kontrahierend ist.

c.) Nachdem Γ abgeschlossen ist, folgt mit dem Kontraktions-
satz von BANACH, daß $\mathcal{K}$ genau einen Fixpunkt in Γ hat
und die Rekursion gegen diesen konvergiert.

Für den Fixpunkt gilt:

$$X = A_{\gamma}^{-1} C + A_{\gamma}^{-1} B X^2$$

also folgt nach Multiplikation mit A_{γ} von links, daß die-
ser Fixpunkt tatsächlich auch Lösung der quadratischen
Gleichung ist.

q.e.d.

Natürlich gilt auch hier wieder eine völlig analoge Aussage für
die Rekursion "von oben nach unten" (vgl. Seite 55/56), bei der
ganz analog zu den Forderungen in Paragraph 7 der Verträglich-
keitsfaktor γ wieder größer gleich 1 sein muß:

(9.2) SATZ

Sei $\gamma \geq 1$, $\|A_\gamma^{-1}C\| < \frac{1}{1+\gamma}$ und $\|A_\gamma^{-1}B\| < \frac{1}{1+\gamma}$

Dann konvergiert die Rekursion

$$X_0 := 0$$

$$X_{i+1} := \gamma A_\gamma^{-1}B + A_\gamma^{-1}C X_i^2 \quad \text{für } i=1,2,3,\dots$$

gegen diejenige Lösung von $CX^2 - A_\gamma X + \gamma B = 0$ mit $\|X\| \leq 1$.

Beweis: analog zu Satz (9.1).

Die entscheidenden Vorteile bei diesen neuen Rekursionen sind folgende:

- * Es muß nicht bei jedem Schritt eine neue Matrix invertiert werden. Die Matrizen $A_\gamma^{-1}B$ und $A_\gamma^{-1}C$ werden nur einmal ganz zu Anfang bereitgestellt und brauchen dann nicht mehr verändert zu werden.
- * Der Programmieraufwand zur Realisierung dieses Verfahrens ist sehr gering. Wenn man für die Berechnung von $A_\gamma^{-1}B$ und $A_\gamma^{-1}C$ fertige Bibliothekssoftware verwendet, wie sie an jedem Rechenzentrum heute vorhanden ist, dann besteht der ganze Algorithmus nur noch aus Matrixmultiplikationen.

Die Konvergenzeigenschaften dieser vereinfachten Rekursion beruhen auf dem gleichen Mechanismus (Fixpunktcontraktion) wie die der Originalrekursion und sind deshalb völlig äquivalent.

Die Überlegenheit der neuen, vereinfachten Rekursion besteht also nur aus einem Gewinn an Rechenzeit. Dieser wird in folgendem Diagramm veranschaulicht:

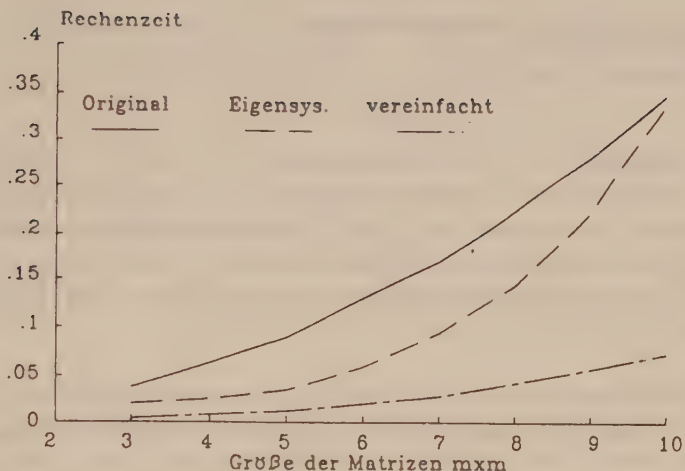


Bild 2: Vergleich der Rechenzeiten

Die Kurven beziehen sich auf:

Original: 10 Iterationen der Originalrekursion (9.1)

vereinfacht: 10 Iterationen der vereinfachten Rekursion gemäß Satz (9.1)

Eigensys.: zum Vergleich die Zeiten, die notwendig sind, um das aufgedoppelte Eigenwertproblem zu lösen und daraus den Fixpunkt zu synthetisieren.
Man erkennt deutlich die höhere Ordnung im Anwachsen der Rechenzeit.

Dem Vergleich liegen die entsprechenden Matrizen einer F.E.-Zerlegung in TRIM3-Elemente an einer 90° -Ecke bei einem einfachen Potentialproblem zugrunde. Die beiden Iterationsverfahren ergaben vergleichbare Genauigkeit. Die Eigenwertzerlegung wurde mit Bibliotheksprogrammen aus der IMSL-Library berechnet.

10) BEDEUTUNG DER LATENTEN WURZELN

Die Eigenwerte λ_j und Eigenvektoren v_j $j=1, \dots, m$ der Fixpunktmatrix X_∞ von $\mathcal{F}$ haben eine ganz konkrete Bedeutung: sie entsprechen den Exponenten s_j bzw. den Winkelfunktionen ψ_j der bereits in der Einleitung angedeuteten Reihenentwicklung. Wir wollen diesen Zusammenhang im folgenden Abschnitt etwas näher untersuchen.

Für feste Werte von N (= Anzahl der verschachtelten Substrukturen) erhält man aus der Rekursionsvorschrift (6.3) (bzw. völlig analog für (6.8)) jeweils eine endliche Folge von Matrizen:

$$X_0^{(N)} \rightarrow X_1^{(N)} \rightarrow X_2^{(N)} \dots \rightarrow X_N^{(N)}$$

Für wachsende N kann man diese zu folgendem Schema anordnen:

k=	N	N-1	N-2	N-3	
N=0	$X_0^{(0)}$				
N=1	$X_1^{(1)}$	$X_0^{(1)}$			
N=2	$X_2^{(2)}$	$X_1^{(2)}$	$X_0^{(2)}$		
N=3	$X_3^{(3)}$	$X_2^{(3)}$	$X_1^{(3)}$	$X_0^{(3)}$	
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
allg. N	$X_N^{(N)}$	$X_{N-1}^{(N)}$	$X_{N-2}^{(N)}$		... $X_0^{(N)}$

Dabei ist für jedes N die Matrix $X_0^{(N)}$ per Definition der Startwert der Iteration. Nach Formel (6.1) also $X_0^{(N)} = K_0^{-1}C$. Die Zeilen des Schemas werden jeweils durch die Rekursionsformel (6.3) von rechts nach links erzeugt. Deshalb sind die Matrizen in den Dia-

gonalen (links oben - rechts unten) alle gleich, d.h.:

$$\begin{aligned} x_k^{(N)} &= x_k^{(N-1)} & \text{f. alle } N \\ & & \text{f. alle } k=0,1,\dots,N-1 \end{aligned} \quad (10.1)$$

Somit steht in jeder der Spalten des Schemas genau die Folge der Fixpunktiterierten zum Operator $\mathcal{F}$.

$$\begin{aligned} x_{N-k}^{(N)} &= \mathcal{F}_{N-k+1}(x_{N-k+1}^{(N-1)}) & \text{f. alle } N=1,2,3,\dots \\ & & \text{f. alle } k=0,\dots,N \end{aligned} \quad (10.2)$$

Nach den Aussagen (und unter den Voraussetzungen) von §7 konvergieren ergo alle Spalten gegen den Fixpunkt X_∞ . Im Grenzübergang $N \rightarrow \infty$ entsteht auf diese Weise eine (unendlich lange) Zeile der Form

$$X_\infty, X_\infty, X_\infty, X_\infty, X_\infty, X_\infty, \dots$$

Wir wollen nun der Einfachheit halber von einem homogenen Problem ausgehen. Dann verschwindet die rechte Seite F und damit auch alle Iterierten H_i in der Rekursion (6.2). Gemäß Formel (6.5) (Seite 56) ergeben sich die Unbekannten U_i rekursiv aus:

$$U_i = -X_{N-i} U_{i-1} \quad i=2,3,\dots,N \quad (10.3)$$

also für $N \rightarrow \infty$:

$$U_i = -X_\infty \cdot U_{i-1} \quad \text{f. alle } i \quad (10.4)$$

und somit:

$$U_{i+1} = (-X_\infty)^i U_1 \quad \text{f. alle } i=1,2,3,\dots \quad (10.5)$$

Zur Erinnerung: die Vektoren U_i waren dabei die Unbekannten jeweils in der i -ten Substruktur, also Näherungswerte für die Lösung an den Knoten $T^i(x_j)$ $j=1,\dots,m$; wobei die x_j die Knoten der 1. Substruktur seien.

Wenn nun die Matrix X_∞ diagonalähnlich ist - eine Voraussetzung, die bereits mehrfach verwendet wurde -, dann läßt sich der Vektor U_1 aus den Eigenvektoren von X_∞ linear kombinieren. Also gilt:

$$U_1 = \sum_{j=1}^m \tilde{k}_j v_j \quad (10.6)$$

Den Vektor der Entwicklungskoeffizienten $\tilde{k} = (\tilde{k}_1, \tilde{k}_2, \dots, \tilde{k}_m)^T$ erhält als Lösung des linearen Gleichungssystems

$$V \tilde{k} = U_1 \quad (10.7)$$

wobei $V = (v_1, \dots, v_m)$ die Matrix der Eigenvektoren ist.

Setzt man (10.6) in (10.5) ein, dann ergibt sich für die Unbekannten in der i-ten Substruktur R_i die Reihenentwicklung

$$\begin{aligned} U_i &= (-X_\infty)^i U_1 = (-X_\infty)^i \sum_{j=1}^m \tilde{k}_j v_j = \\ &= \sum_{j=1}^m (-\lambda_j)^i \tilde{k}_j v_j \end{aligned} \quad (10.8)$$

Wir wollen nun eine entsprechende Formel für die wahre Lösung herleiten. Dazu betrachten wir ein Problem, bei dem die wahre Lösung eine Reihenentwicklung folgender Gestalt hat:

$$u(r, \omega) = \sum_{j=1}^{\infty} k_j r^{s_j} f_j(\omega) \quad (10.9)$$

r, ω sind dabei Polarkoordinaten (ω kann auch ein Vektor sein!). Eine solche Entwicklung gilt z.B. im allgemeinen bei homogenen Potentialproblemen an einspringenden Ecken (vgl. LEHMANN (1957), KONDRATJEV (1967), KELLOG (1971), STRANG, FIX (1973)) oder bei Interfaceproblemen (vgl. KELLOG (1971), (1975), STRANG, FIX (1973)). Die SNEDDON'sche Formel für die Verschiebungen, die man in der Bruchmechanik allgemein verwendet, (siehe z.B. IRWIN (1958)) ist ebenfalls von dieser Form.

Als affine Transformation T für die infinite Substrukturierung kommt dann sinnvollerweise eine zentrische Streckung/Stauchung um einen Faktor t zur Anwendung.

Wenn $(r_1, \omega_1) = x_1$ $l=1, \dots, m$ die Polarkoordinaten der Knoten in der 1. Substruktur sind, dann folgt für die wahren Werte der Lösung an den Knoten $T^i(x_1) = (t^i r_1, \omega_1)$ eine Darstellung als:

$$\begin{pmatrix} u(t^i r_1, \omega_1) \\ u(t^i r_2, \omega_2) \\ \vdots \\ u(t^i r_m, \omega_m) \end{pmatrix} = t^{i s_1} k_1 \begin{pmatrix} f_1(\omega_1) r_1^{s_1} \\ f_1(\omega_2) r_2^{s_1} \\ \vdots \\ f_1(\omega_m) r_m^{s_1} \end{pmatrix} + t^{i s_2} k_2 \begin{pmatrix} f_2(\omega_1) r_1^{s_2} \\ f_2(\omega_2) r_2^{s_2} \\ \vdots \\ f_2(\omega_m) r_m^{s_2} \end{pmatrix} + \dots \quad (10.10)$$

Vergleicht man nun die Reihe für die wahre Lösung (10.10) mit der für die Näherung (10.8), so kann man folgern, daß die Terme der einzelnen i -Potenzen einander gliedweise entsprechen müssen.

Also für den ersten Summanden:

$$t^{i s_1} k_1 \begin{pmatrix} f_1(\omega_1) r_1^{s_1} \\ f_1(\omega_2) r_2^{s_1} \\ \vdots \\ f_1(\omega_m) r_m^{s_1} \end{pmatrix} = (-\lambda_1)^i \tilde{k}_1 v_1 + \underline{\epsilon} \quad (10.11)$$

mit einem (hoffentlich sehr kleinen) Fehlervektor $\underline{\epsilon}$. Da der Eigenvektor v_1 nur bis auf einen Faktor eindeutig bestimmt ist,

können wir sie so skalieren, daß gilt:

$$t^{is_i} = (-\lambda_1)^i + \varepsilon^* \quad (10.12)$$

und für die Vektoren:

$$v_1 = k_1 / \tilde{k}_1 \cdot \begin{pmatrix} f_1(\omega_1) r_1^{s_1} \\ \vdots \\ f_1(\omega_m) r_m^{s_m} \end{pmatrix} + \hat{\varepsilon} \quad (10.13)$$

Ganz analoge Formeln gelten natürlich auch für die höheren Eigenwerte / Eigenvektoren - allerdings mit entsprechend größeren Abweichungen.

Aus Gleichung (10.12) bzw. den analogen Gleichungen für höhere Indices i lassen sich sofort die Singularitätenordnungen s_i - zumindest näherungsweise - berechnen:

$$s_i = \log(-\lambda_i) / \log(t) \quad (10.14)$$

Normalerweise ergeben sich die Eigenwerte λ_i als negative Zahlen, die betragsmäßig kleiner als 1 sind. Für zentrische Stauungen, also mit $t < 1$, werden die s_i folglich positiv.

Zu beachten ist, daß die Formel auch im Komplexen gilt. Die Eigenwerte λ_i können nämlich durchaus auch komplex werden. In der Tat zeigen sich in numerischen Experimenten meist höhere latente Wurzeln, die komplex sind. Mit

$$\begin{aligned} \lambda &= \xi + i\eta \\ s &= a + ib \end{aligned}$$

folgen dann mit (10.14) in der Form

$$\begin{aligned} a &= \log(\sqrt{\xi^2 + \eta^2}) / \log(t) \\ b &= (\arctan(\eta/\xi) + 2k\pi) / \log(t) \end{aligned} \quad (10.15)$$

($k \in \mathbb{N}$ entsprechend dem Hauptzweig des komplexen Logarithmus)
auch komplexe Singularitätenordnungen. Das folgende Lemma gibt ein Kriterium für ihr Auftreten:

(10.1) LEMMA:

Sei $0 \neq \gamma \in \mathbb{R}$, $\lambda \in \mathbb{C}$, $v \in \mathbb{C}^m$ mit $\lambda^2 \gamma Bv - \lambda Av + B^T v = 0$.

Sei A symmetrisch und es gelte $(v^T A v)^2 > 4\gamma (v^T Bv)(v^T B^T v)$

Dann gilt:

$$\lambda \in \mathbb{R} \quad \Leftrightarrow \quad v^T Bv \in \mathbb{R}$$

Beweis:

Multipliziert man $\lambda^2 \gamma Bv - \lambda Av + B^T v = 0$ von links mit v^T , so folgt:

$$\lambda^2 \gamma v^T Bv - \lambda v^T A v + v^T B^T v = 0$$

Wenn $v^T Bv = 0$ ist, so folgt $v^T B^T v = 0$ und somit $\lambda v^T A v = 0$
also $\lambda = 0$.

Ansonsten gilt:

$$\lambda_{1|2} = \frac{v^T A v \pm \sqrt{(v^T A v)^2 - 4\gamma (v^T Bv)(v^T B^T v)}}{2\gamma v^T Bv}$$

Aus der Symmetrie von A folgt, $v^T A v \in \mathbb{R}$. Nach Voraussetzung ist der Term unter der Wurzel nicht negativ, die Wurzel wird also nicht komplex. Damit ergibt sich die Behauptung.

q.e.d.

In vielen Fällen handelt es sich bei diesen komplexen Singularitätenordnungen wohl um numerische "Schmutzeffekte". Ihr Auftreten läßt sich allerdings nicht prinzipiell ausschließen, da durchaus physikalisch sinnvolle Problemstellungen tatsächlich auf komplexe Singularitätenordnungen führen können, sodaß die Berechnungsmethode die Möglichkeit des Komplexwerdens stets beinhalten muß.

Beispielsweise immer dann, wenn eine Riß an der Grenzfläche zwischen 2 verschiedenen Materialien endet, entstehen derartige Singularitäten (siehe etwa ENGLAND (1965), BOGY (1968)). Die physikalische Interpretation dieses Phänomens ist sehr problematisch. In der Arbeit von LÖBEL, ZEILINGER und DESKA (1984) wird der Themenkreis unter dem Aspekt der Anwendung der Bruchmechanik auf moderne Faserverbundwerkstoffe diskutiert. Letzten Endes muß man wohl davon ausgehen, daß in diesen Fällen die mathematische Differentialgleichung keine vernünftige Beschreibung der Wirklichkeit mehr ist.

Bei der Anwendung des Verfahrens auf unbeschränkte Gebiete gelten natürlich ganz analoge Zusammenhänge. Da dabei aber eine zentrische Streckung mit einem Faktor $t > 1$ zur Anwendung kommt, werden die s_i gemäß Formel (10.14) negativ. Dies entspricht genau dem Abklingverhalten der wahren Lösung. Die Eigenvektoren der Fixpunktmatrix Y_∞ von $\mathcal{G}$ entsprechen auch hier im wesentlichen der Winkelverteilung.

KAPITEL 4 : KONVERGENZ UND NUMERISCHE ERGEBNISSE

11) KONVERGENZ BEI SINGULARITÄTEN UND UNBESCHRÄNKTEN GEBIETEN

In diesem Abschnitt soll gezeigt werden, daß das vorgeschlagene Verfahren der infiniten Substrukturkondensation die optimale Konvergenzordnung der Finite-Elemente-Methode auch bei Singularitäten und unbeschränkten Gebieten erhalten kann.

Dabei wollen wir zunächst davon ausgehen, daß für die wahre Lösung u eine Darstellung in der Form

$$u(r, \omega) = \sum_{i=1}^{\infty} k_i r^{s_i} \psi_i(\omega) \quad (11.1)$$

mit $s_1 < s_2 < s_3 < \dots$ und $0 < s_1 < 1$ gilt. Dabei seien $r=r(x)$, $\omega=\omega(x)$ für $x \in \mathbb{R}^n$ wieder n -dimensionale Polarkoordinaten $r \in \mathbb{R}^+$, $\omega \in \mathbb{R}^{n-1}$. Eine solche Entwicklung gilt beispielsweise bei einspringenden Ecken, Interfaceproblemen mit Ecken oder sprunghaft sich ändernden Randbedingungen (siehe etwa KONDRATJEV (1967), KELLOG (1971), (1975), STRANG, FIX (1973), LEHMANN (1957), FENNER (1976)).

Für eine Folge von Näherungslösungen u_h aus verschiedenen Räumen V_h $h = h_1, h_2, \dots$ wird der Fehler $u - u_h$ normalerweise durch eine Abschätzung der Form

$$\|u - u_h\| \leq C h^{k+1} \quad (11.2)$$

beschränkt, die asymptotisch für kleine h gilt. Dabei ist der Parameter h eine charakteristische Größe der jeweiligen F.E.-Zerlegung des Raumes V_h - zum Beispiel der maximale Elementdurchmesser.

Mit dem Buchstaben C werden wir in diesem Abschnitt allgemein

Konstanten $C > 0$ bezeichnen, die nicht näher spezifiziert werden und an verschiedenen Stellen nicht unbedingt gleich zu sein brauchen.

Der Exponent k in (11.2), genannt Konvergenzordnung, hängt dabei im wesentlichen ab von der Glattheit der wahren Lösung und den Approximationseigenschaften der Räume V_h an Funktionen aus V .

Das Bindeglied zwischen den generellen Approximationseigenschaften von V_h und der speziellen F.E.-Näherung u_h ist das CEA'sche Lemma, das wir hier aus CIARLET (1978) zitieren:

(11.1) LEMMA (CEA)

Es gibt eine Konstante $C > 0$ unabhängig vom Unterraum V_h so, daß

$$\| u - u_h \| \leq C \inf_{v_h \in V_h} \| u - v_h \|$$

Beweis: siehe CIARLET (1978), Seite 104, Theorem 2.4.1 .

Demnach genügt es also für Fehlerabschätzungen jeweils die Interpolierenden aus dem Teilraum V_h an ein $u \in V$ näher zu untersuchen.

Der Interpolationsoperator $\Pi_{\Omega, h}$ zu einer F.E.-Zerlegung zum Parameter h eines Gebietes Ω sei definiert als:

$$\begin{aligned} \Pi_{\Omega, h}: V &\longrightarrow V_h \\ u &\longrightarrow \sum_{i=1}^m \phi_i(u) \psi_i \end{aligned} \quad (11.3)$$

wobei : $\{\psi_1, \psi_2, \dots, \psi_m\}$ Basis von V_h

$\phi_i \quad i=1, \dots, m$ Funktional des zu ψ_i gehörigen Freiheitsgrades (vgl. §1)

Die Aussage in CEA's Lemma (11.1) bezieht sich dabei auf den Grundraum V und seine Norm (vgl. §1), im allgemeinen also eine SOBOLEV-Norm. Wir wollen diese wie folgt notieren:

(11.2) DEFINITION:

Sei $\Omega \subset \mathbb{R}^n$ beschränktes Gebiet mit LIPSCHITZ-stetigem Rand. $L_p(\Omega)$ $1 < p < \infty$ sei der Raum der Äquivalenzklassen von meßbaren Funktionen auf Ω , deren p -te Potenzen integrierbar sind. Für ein $u \in L_p(\Omega)$ sei:

$$\|u\|_{0,p} := \|u\|_{0,p,\Omega} := \left(\int_{\Omega} (|u(x)|^p) dx \right)^{1/p}.$$

$L_{\infty}(\Omega)$ sei der Raum der meßbaren, wesentlich beschränkten Funktionen über Ω , mit der Norm:

$$\|u\|_{0,\infty} := \|u\|_{0,\infty,\Omega} := \operatorname{ess\,sup}_{x \in \Omega} |u(x)|.$$

Mit $W^{s,p}(\Omega)$ $1 < p < \infty$, $s \in \mathbb{N}_0$ bezeichnen wir den SOBOLEV-Raum der Funktionen aus $L_p(\Omega)$, die verallgemeinerte Ableitungen bis zur Ordnung s in $L_p(\Omega)$ besitzen. Für $u \in W^{s,p}(\Omega)$ sei

$$|u|_{s,p} := |u|_{s,p,\Omega} := \left(\sum_{|\alpha|=s} \|D^{\alpha} u\|_{0,p}^p \right)^{1/p}$$

und:

$$\|u\|_{s,p} := \|u\|_{s,p,\Omega} := \left(\sum_{|\alpha| \leq s} \|D^{\alpha} u\|_{0,p}^p \right)^{1/p}$$

Als Abkürzung verwenden wir auch die Räume $H^s(\Omega) := W^{s,2}(\Omega)$ $s \in \mathbb{N}_0$ mit zugehöriger Norm.

Für eine detaillierte Beschreibung dieser Räume und ihrer Eigenschaften sei z.B. auf ADAMS (1975) verwiesen.

Für hinreichend glatte Funktionen $w \in W^{k+1,p}(\Omega)$, lassen sich für die meisten finiten Elemente (E, P, Σ) Abschätzungen des Interpolations-

tionsfehlers in folgender Form zeigen:

$$\|w - \prod_{E,h} w\|_{m,q,E} \leq C |E|^{1/q-1/p} h^{k+1-m} |w|_{k+1,p,E} \quad (11.4)$$

Dabei ist $|E| := \int_E 1 \, dx$ das Volumen des Elements E , h der größte Elementdurchmesser. Die möglichen Werte für m, q, p werden durch bestimmte Inklusionsbedingungen begrenzt (CIARLET (1978), Seite 124, Theorem 3.1.6).

Im Prinzip findet man diese Abschätzungen, indem man w elementeweise in eine TAYLORreihe entwickelt und das Interpolationspolynom durch Koeffizientenabgleich bestimmt. Die Ordnung $k+1$ ist dann offenbar gerade die Ordnung des TAYLORSchen Restgliedes – also um eines mehr als der Polynomgrad der verwendeten Ansatzfunktionen (bei vollständigen Polynomen). Durch Aufsummieren der Fehler in den einzelnen Elementen gelangt man dann zu (11.4).

Mit dem CEA'schen Lemma folgt dann unmittelbar eine Fehlerabschätzung wie in (11.1). Da eine bessere Konvergenzordnung im allgemeinen nicht erwartet werden kann, nennt man dies auch die optimale Konvergenzordnung.

Eine Funktion der Gestalt wie in Formel (11.1) ist in Gebieten, die den kritischen Punkt $r=0$ enthalten aber nicht glatt. Ihre Ableitungen wachsen bei $r=0$ über alle Grenzen. Eine genauere Analyse zeigt, daß sie nur noch in $H^{1+s-\varepsilon}(\Omega)$ $\varepsilon > 0$ ist. (vgl. SCHATZ, WAHLBIN (1978), STRANG, FIX (1973)). Auf die Theorie der SOBOLEV-Räume zu reellen Exponenten wollen wir hier nicht näher eingehen (siehe hierzu z.B. LIONS, MAGENES (1968)) und uns nur mit der Feststellung begnügen, daß

$$H^{i+1}(\Omega) \subset H^s(\Omega) \subset H^i(\Omega) \quad \text{für } i \in \mathbb{N}, s \in \mathbb{R} \text{ und } i \leq s \leq i+1.$$

Zum Beispiel ist an der Spitze eines Risses die Verschiebung von der Form (11.1) mit $s_1 = .5$. Sie ist also nur noch in $H^1(\Omega)$, aber nicht in $H^2(\Omega)$. Damit sind die üblichen Abschätzungen nicht mehr möglich und in der Tat zeigt sich, daß die F.E.-Methode an

derartigen Singularitäten bei gleichmäßiger Elemententeilung nicht die optimale Konvergenzordnung erreicht.

In diesem Fall ist es sinnvoll die Elemententeilung lokal an der Singularität zu verfeinern. Denn außer in einer Umgebung der Singularität ist die Lösung ja harmlos (d.h. aus C^∞).

Dazu folgende Aussage:

(11.3) LEMMA:

Sei $\Omega \subset \mathbb{R}^n$ beschränkt. $\rho := \inf_{x \in \Omega} r(x) > 0$.

$v: \Omega \rightarrow \mathbb{R}$ sei von der Form wie Gl. (11.1) mit $\psi_i \in C^\infty(\mathbb{R}^{n-1})$.

Dann gilt für alle $k \in \mathbb{N}$, $1 \leq p < \infty$, $\alpha \in \mathbb{N}^n$

$$v \in W^{k,p}(\Omega) \quad \text{und} \quad \|v\|_{k,p,\Omega} \leq C \rho^{s_\alpha - k} |\Omega|^{1/p}$$

Beweis:

Durch gliedweise Differentiation und Induktion nach α läßt sich zeigen, daß

$$\partial^\alpha v = \sum_{i=1}^{\infty} k_i^{(\alpha)} r^{s_i - |\alpha|} \psi_i^{(\alpha)}(\omega)$$

mit neuen Koeffizienten $k_i^{(\alpha)}$ und Winkelfunktionen $\psi_i^{(\alpha)}$, die o.B.d.A. auf 1 normiert werden können.

Es folgt $|\partial^\alpha v| = O(r^{s_\alpha - |\alpha|})$ für $r \rightarrow 0$ und damit:

$$\begin{aligned} \|v\|_{k,p,\Omega} &= \left(\int_{\Omega} \left| \sum_{|\alpha|=k} \partial^\alpha v \right|^p dx \right)^{1/p} \leq \\ &\leq \left(\int_{\Omega} \left(\sum_{|\alpha|=k} |\partial^\alpha v| \right)^p dx \right)^{1/p} \leq \\ &\leq C \rho^{s_\alpha - k} \left(\int_{\Omega} 1 dx \right)^{1/p} = \\ &= C \rho^{s_\alpha - k} |\Omega|^{1/p} \end{aligned}$$

q.e.d.

Das Entscheidende bei einer lokalen Verfeinerung ist, daß die Elementgröße in Richtung auf die Singularität zu schneller verschwindet, als der Interpolationsfehler durch die steilen Gradienten anwächst. SCHATZ und WAHLBIN (1978), (1979) haben gezeigt, daß eine endliche Verfeinerung (mit bestimmten Bedingungen) ausreichend ist, um die optimale Konvergenzordnung zu erhalten. Sie konnten dies sogar und insbesondere für die H^1 -Norm zeigen.

Allerdings ist die Erzeugung solcher Verfeinerungen in der Praxis recht mühselig und durch die größere Anzahl von Freiheitsgraden, die dabei entstehen, auch sehr rechenaufwendig.

Die Vermutung, daß die infinite Substrukturierung, die ja eine unendliche Verfeinerung simuliert, ebenfalls die optimale Konvergenzordnung erhält ist also naheliegend. THATCHER (1975) beweist für eine spezielle Zerlegung beim MOTZ-Problem (MOTZ (1946)), die Abschätzung

$$\|u - u_h\|_1 \leq C h$$

die bei den verwendeten TRIM3-Elementen (=linear, also $k=1$, $m=1$) optimal ist.

HOUDE HAN (1982) benützt gewichtete SOBOLEV-Normen und beweist damit - ebenfalls nur für lineare Elemente - optimale Konvergenz bei der Anwendung auf 2-dimensionale Interfaceprobleme mit Ecken.

Im Folgenden wollen wir diese Aussagen an unsere Form der infiniten Substrukturierung anpassen und damit auf andere Normen, Elemente und Dimensionen verallgemeinern. Der Grundgedanke besteht - wie bei THATCHER - darin, die Standardabschätzung, die man für die 1. Substruktur hat, entsprechend auf alle transformierten Wiederholungen zu übertragen und die einzelnen Fehler aufzusummieren. Es entsteht dabei eine geometrische Reihe, die summierbar ist.

(11.4) SATZ:

Sei $\Omega_0 \subset \mathbb{R}^n$, $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $(r, \omega) \rightarrow (tr, \omega)$ zentrische Stauchung mit $t < 1$.

$\Omega := \bigcup_{i=0}^{\infty} T^i(\Omega_0)$ infinite Substrukturierung.

Ω_0 sei regulär in affin äquivalente Elemente mit maximalem Durchmesser h unterteilt, sodaß

- a) die Zerlegung mit T kompatibel ist - d.h. durch T eine gültige F.E.-Zerlegung von Ω erzeugt wird. (vgl. §5)
- b) für alle $w \in W^{k+1,p}(\Omega_0)$ eine Abschätzung des Interpolationsfehlers auf Ω_0 gemäß Gl. (11.5) gilt.

Die Funktion $v: \Omega \rightarrow \mathbb{R}$ sei von der Form gemäß Gl. (11.1).

Dann gilt für alle $m < \min(n/q + s_1, k+1)$

$$\|v - \Pi_{\Omega,h} v\|_{m,q,\Omega} \leq C \cdot \frac{1}{1-t^{(n+s_1q-mq)}} \cdot h^{k+1-m}$$

Beweis:

Es sei $\Omega_i := T^i(\Omega_0)$, h_i max. Durchmesser von Ω_i und $\rho_i := \inf_{x \in \Omega_i} r(x)$ für $i=1,2,3,\dots$

Dann ist :

$$\|v - \Pi_{\Omega,h} v\|_{m,q,\Omega}^q = \sum_{i=0}^{\infty} \|v - \Pi_{\Omega_i,h_i} v\|_{m,q,\Omega_i}^q \leq (*)$$

aufgrund der Affinäquivalenz der einzelnen Substrukturen gilt für jedes Ω_i wieder eine Abschätzung des Interpolationsfehlers gemäß (11.5). Die jeweiligen Konstanten C sind dabei alle gleich (CIARLET (1978), SEITE 124, Theorem 3.1.6.). Damit:

$$(*) \leq \sum_{i=0}^{\infty} (C / \rho_i)^{1/q-1/p} h_i^{k+1-m} |v|_{k+1,p,\Omega_i}^q \leq$$

nach Lemma (11.3):

$$\leq C \sum_{i=0}^{\infty} (|\Omega_i|^{1/q} h_i^{k+1-m} \rho_i^{s_1-k-1})^q \quad (**)$$

Die Größen $|\Omega_i|$, h_i , ρ_i $i=1,2,3,\dots$ lassen sich dabei jeweils durch $|\Omega_0|$, h , ρ ausdrücken:

$$\begin{aligned} |\Omega_i| &= t^{ni} |\Omega_0| & (\text{Volumen}) \\ h_i &= t^i h & (\text{max. Durchmesser}) \\ \rho_i &= t^i \rho & (\text{Innenradius}) \end{aligned}$$

$$(**) \leq (C |\Omega_0|^{1/q} h^{k+1-m} \rho^{s_1-k-1} \sum_{i=0}^{\infty} t^{i(n/q+s_1-m)})^q$$

wegen $t < 1$ und $n/q + s_1 - m > 0$ nach Voraussetzung ist diese Reihe summierbar und es ergibt sich die Behauptung.

q.e.d.

Wählt man anstelle der zentrischen Stauchung mit $t < 1$ eine zentrische Streckung, also $t > 1$, so wird durch die infinite Substrukturierung

$$\Omega = \bigcup_{i=0}^{\infty} T^i(\Omega_0)$$

ein unbeschränktes Gebiet Ω in finite Element zerlegt. Dies ermöglicht die Analyse von Außenraumproblemen oder Problemen auf unbeschränkten Gebieten mit der Methode der Finiten Elemente.

Es gibt dabei eine ganze Reihe von Parallelen zu Problemen mit Singularitäten. Die wahre Lösung ist hier allerdings normalerweise völlig glatt (wenn keine Inhomogenitäten vorliegen) und klingt zusammen mit allen Ableitungen für $r \rightarrow \infty$ ab.

Wir wollen davon ausgehen, daß auch hier eine Entwicklung der Form

$$u(r, \omega) = \sum_{i=1}^{\infty} k_i r^{s_i} \psi_i(\omega) \quad (11.5a)$$

gilt, wobei uns hier große r besonders interessieren und

$$0 > s_1 \geq s_2 \geq s_3 \geq \dots \quad (11.5b)$$

gelte. GOLDSTEIN (1981) zeigt die Gültigkeit für Außenraumprobleme im R^3 .

Je nach Formulierung des Problems wird das Abklingverhalten der Lösung bei Unendlich explizit als zusätzliche Randbedingung gefordert. (siehe z.B. LEIS (1967)). Es ist daher sinnvoll und möglich für wachsendes r die Elemente zu vergrößern und dennoch gleiche Approximationsgenauigkeit zu behalten.

GOLDSTEIN (1981) untersucht Außenraumprobleme im R^3 mit der F.E.M. Er schneidet dabei das unbeschränkte Gesamtgebiet bei großen Radien $r=R_i$ ab und stellt auf dem so geschaffenen künstlichen Rand eine homogene Dirichletbedingung als Ersatz für die Abklingbedingung im Unendlichen. Diese Ersatzprobleme werden dann mit einem F.E.-Netz mit nach außen gröber werdenden Elementen gelöst. Es zeigt sich, daß es auf diesem Weg möglich ist, die optimale Konvergenzordnung zu erhalten.

Allerdings ist hier der maximale Elementdurchmesser keine sehr vernünftige Bezugsgröße mehr. BABUSKA (1972) führt daher für derartige Probleme eine mittlere Länge H ein, durch:

$$H := (\dim V_h)^{-1/n} \quad (11.6)$$

Es genügt also offenbar hier bereits eine endliche Vergrößerung in Analogie zur endlichen Verfeinerung an Singularitäten zum Erhalt der optimalen Konvergenzordnung.

Die infinite Substrukturierung, die ja eine unendliche Vergrößerung simulieren kann, sollte also auch hier imstande sein die optimale Konvergenz zu erhalten.

Der folgende Satz, der im wesentlichen völlig analog zu Satz (11.4) ist, schätzt dazu den Interpolationsfehler in einem unendlichen Gebiet ab. Als charakteristische Länge der F.E.-Zerlegung wird dabei der größte Elementdurchmesser der innersten Substruktur verwendet, da die Raumdimension $\dim(V_h)$ für (11.6) ja nicht endlich bleibt.

(11.5) SATZ:

Sei $\Omega_0 \subset \mathbb{R}^n$, $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $(r, \omega) \rightarrow (tr, \omega)$ zentrische Streckung mit $t > 1$.

$\Omega := \bigcup_{i=0}^{\infty} T^i(\Omega_0)$ infinite Substrukturierung.

Ω_0 sei regulär in affin äquivalente Elemente mit maximalem Durchmesser h unterteilt, sodaß

- a) die Zerlegung mit T kompatibel ist - d.h. durch T eine gültige F.E.-Zerlegung von Ω erzeugt wird. (vgl. §5)
- b) für alle $w \in W^{k+1,p}(\Omega_0)$ eine Abschätzung des Interpolationsfehlers auf Ω_0 gemäß Gl. (11.5) gilt.

Die Funktion $v: \Omega \rightarrow \mathbb{R}$ sei von der Form gemäß Gl. (11.5).

Dann gilt für alle $n/q + s_1 < m < k+1$

$$\|v - \Pi_{\Omega, h} v\|_{m, q, \Omega} \leq C \frac{1}{1-t^{(n+s_1 q - m q)}} h^{k+1-m}$$

Beweis:

Der Beweis ist weitgehend analog zum Beweis von Satz (11.4). Die Aussage von Lemma (11.3) gilt auch in diesem Fall analog. Man erhält formal die gleiche Summe, die nun

allerdings wegen $t > 1$ und $(n/q + s_1 - m) < 0$ summierbar ist.

q.e.d.

Die Randbedingung im Unendlichen, die durch die infinite Substrukturierung angenommen wird, kann man sich folgendermaßen plausibel machen:

Würde man eine endliche, N-fache Gittervergrößerung vornehmen und auf dem äußeren Rand dann eine homogene Dirichletbedingung ($u=0$) stellen, dann kann dies als wesentliche Randbedingung in die Steifigkeitsmatrix aufgenommen werden (vgl. Seite 57). Es wird dabei nur die Matrix A_0 verändert. Da sich dies aber nur in der Startmatrix für die Rekursion (6.8) bemerkbar macht, ändert sich nichts am Fixpunkt.

Die infinite Substrukturierung bei unbeschränktem Gebiet simuliert also das Abklingverhalten ähnlich wie der GOLDSTEIN'sche Ansatz (GOLDSTEIN (1981)), als Grenzübergang einer homogenen Dirichletbedingung auf immer größeren Gebieten.

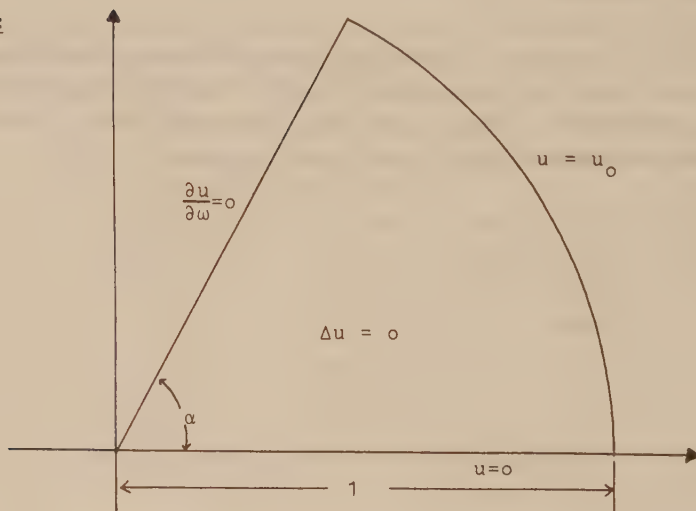
12) NUMERISCHE ERGEBNISSE

Zur Verifizierung der erhaltenen Aussagen wurden verschiedene numerische Untersuchungen gemacht. Um die Brauchbarkeit von Formel (10.14) zur numerischen Bestimmung der Singularitätenordnung (bzw. des Abklingverhaltens) zu testen, wurde das folgende Problem analysiert:

$$\Omega := \{ (r, \omega) / 0 \leq r \leq 1, 0 \leq \omega \leq \alpha \} \subset \mathbb{R}^2$$

$$\begin{aligned} \Delta u &= 0 && \text{in } \Omega \\ u(r, 0) &= 0 && 0 \leq r \leq 1 \\ \frac{\partial u}{\partial \omega} u(r, \alpha) &= 0 && 0 \leq r \leq 1 \\ u(1, \omega) &= u_0 && 0 \leq \omega \leq \alpha \end{aligned}$$

Skizze:



Mit dem Ansatz $u(r, \omega) = r^s g(\omega)$ findet man das Eigenwertproblem (vgl. KELLOG (1975)):

$$s^2 g(\omega) + g''(\omega) = 0$$

mit den Randbedingungen

$$\begin{aligned} g(0) &= 0 \\ g'(\alpha) &= 0 \end{aligned}$$

Die allgemeine Lösung lautet

$$g(\omega) = A \sin(s\omega) + B \cos(s\omega)$$

Damit auch die Randbedingungen erfüllt sind, müssen die Konstanten A, B das folgende lineare Gleichungssystem erfüllen:

$$\begin{pmatrix} 0 & 1 \\ s \cos(s\alpha) & -s \sin(s\alpha) \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Nichttriviale Lösungen kann es nur geben, wenn die Determinante dieses Systems verschwindet, also

$$s=0 \text{ oder } s = \frac{2n-1}{2} \frac{\pi}{\alpha} \quad n=1,2,3,\dots$$

gilt. Die wahre Lösung ergibt sich dann durch Linearkombination aller dieser Eigenlösungen in der Form:

$$u(r, \omega) = \sum_{i=0}^{\infty} k_i r^{s_i} g_i(\omega)$$

mit den Singularitätenordnungen $s_i = \frac{2n-1}{2} \frac{\pi}{\alpha}$, $s_0=0$. Die Koeffizienten k_i sind so zu bestimmen, daß $u(1, \omega) = u_0$ wird.

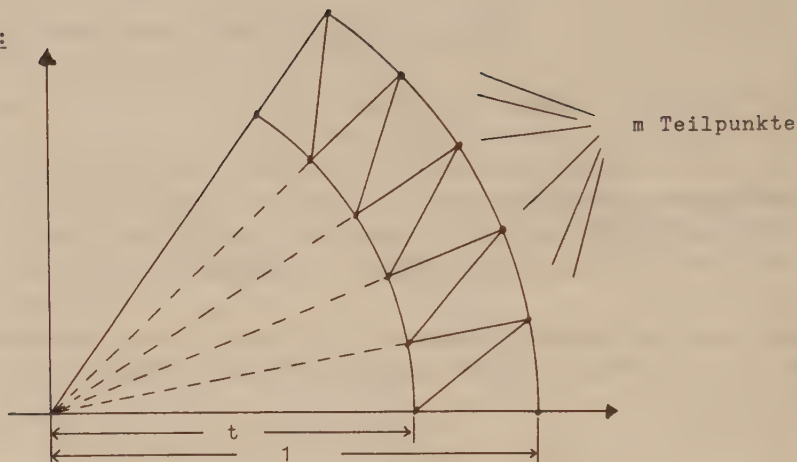
Die numerische Lösung des Problems wurde durch infinite Substrukturierung folgender Bauart gewonnen:

Als affine Abbildung eine zentrische Stauchung

$$T(r, \omega) = (tr, \omega) \quad \text{mit } 0 < t < 1$$

und als 1. Substruktur den Ringsektor $R_1 := \{(r, \omega) / t \leq r \leq 1, 0 \leq \omega \leq \alpha\}$ zerlegt in lineare Dreieckselemente (TRIM3).

Skizze:



Die Steifigkeitsmatrix zu dieser 1. Substruktur ist sehr einfach zu berechnen. Die zugrundeliegende Bilinearform ist invariant $\gamma=1$ unter Drehungen und zentrischen Streckungen (siehe Seite 38), damit genügt es, die Elementsteifigkeitsmatrix für ein Elementpaar im gleichen Sektor bereitzustellen. Für alle anderen Sektoren sind es dann die gleichen.

Die Matrizen A , $B = C^T$ die sich aus dieser Zerlegung ergeben, haben jeweils die Größe $m \times m$. Die latenten Wurzeln wurden durch Lösen des verdoppelten Eigenwertproblems der Form (9.3) bestimmt. Dazu wurden entsprechende Bibliotheksprogramme aus der IMSL-Library des LEIBNIZ-RECHENZENTRUMS München verwendet.

Die folgende Liste gibt die Eigenwerte und die sich daraus gemäß Formel (10.14) ergebenden Singularitätenordnungen für den Fall $m=10$, $\alpha = \pi$, und $t=.8$. Man beachte, daß die Struktur der Eigenwerte tatsächlich der auf Seite 79 vorhergesagten entspricht.

Bei diesem Winkel $\alpha=\pi$ wird im wesentlichen die Singularitätenverteilung des MOTZ-Problems (MOTZ (1946), THATCHER (1975)) beschrieben. Die wahren Werte wären

$$s = 1/2, 3/2, 5/2, 7/2, \dots$$

ZAHL DER RANDKNOTEN: 10
 STRECKUNGSFAKTOR : .800
 ECKWINKEL ALFA : 180.00000

ROUTINE QUADEIG:
 FEHLERCODE: 0

**** EIGENWERTE ****

NR	REAL	IMAG	ITER	SING.-ORD
1	***INDEF***	***INDEF***	0	
2	0.000000	0.000000	0	***INDEF***
3	-3.346782	0.000000	0	-5.41355 0.00000
4	-3.238291	0.000000	0	-5.26587 0.00000
5	-3.032221	0.000000	3	-4.97122 0.00000
6	-2.748863	0.000000	3	-4.53156 0.00000
7	-2.415009	0.000000	2	-3.95128 0.00000
8	-2.060111	0.000000	2	-3.23899 0.00000
9	-1.712072	0.000000	2	-2.40968 0.00000
10	-1.393415	0.000000	2	-1.48675 0.00000
11	-1.118705	0.000000	2	-.50269 0.00000
12	-.893891	0.000000	2	.50269 0.00000
13	-.717661	0.000000	2	1.48675 0.00000
14	-.584087	0.000000	2	2.40968 0.00000
15	-.485411	0.000000	2	3.23899 0.00000
16	-.414077	0.000000	2	3.95128 0.00000
17	-.363787	0.000000	2	4.53156 0.00000
18	-.329791	0.000000	2	4.97122 0.00000
19	-.298795	0.000000	2	5.41355 0.00000
20	-.308805	0.000000	4	5.26587 0.00000

Tabelle 1: Eigenwerte und daraus berechnete Singularitätenordnungen beim MOTZ-Problem für $m=10$, $t=.8$.

In den Experimenten hat sich gezeigt, daß die Genauigkeit der numerischen Singularitätenordnung sehr stark vom Streckungsfaktor t abhängt. Die Anzahl m der Teilpunkte geht erst in 2. Linie ein.

Die Graphiken auf den folgenden Seiten veranschaulichen die Konvergenz der aus den 1. und 2. Eigenwerten numerisch bestimmten Ordnungen gegen die wahren Werte $1/2$ bzw. $3/2$. Dazu muß noch erwähnt werden, daß die Konvergenz der rekursiven Verfahren zur Bestimmung der Grenzmatrix X , wie in Kapitel 9 beschrieben, für große Werte von t immer langsamer wird. Es wird einem also nichts geschenkt - die hohe Genauigkeit bei $t \approx 1$ muß mit einem hohen Aufwand erkauft werden.

Singularitätenordnung
numerische Werte zu $s = .5$

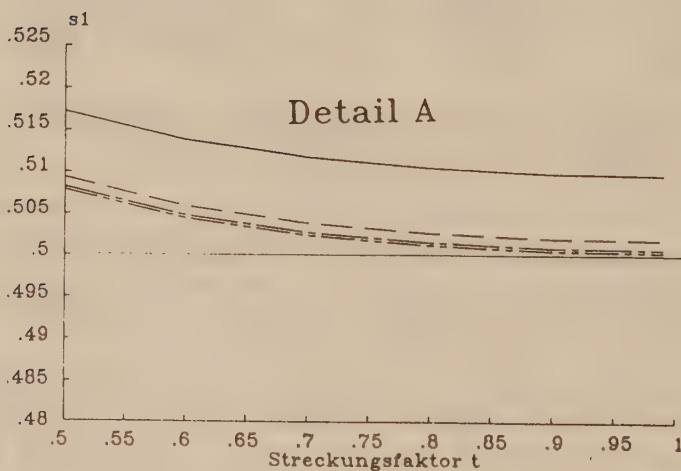
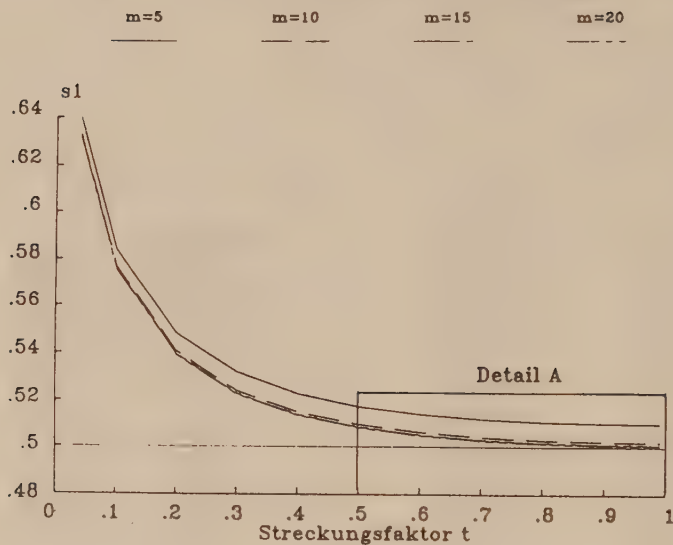


Bild 3: Konvergenz der numerischen Werte der Singularitätenordnungen beim MOTZ-Problem für verschiedene Streckungsfaktoren t . Der wahre Wert ist $s_1 = .5$.

Singularitätenordnung numerische Werte zu $s = 1.5$

$m=5$

$m=10$

$m=15$

$m=20$

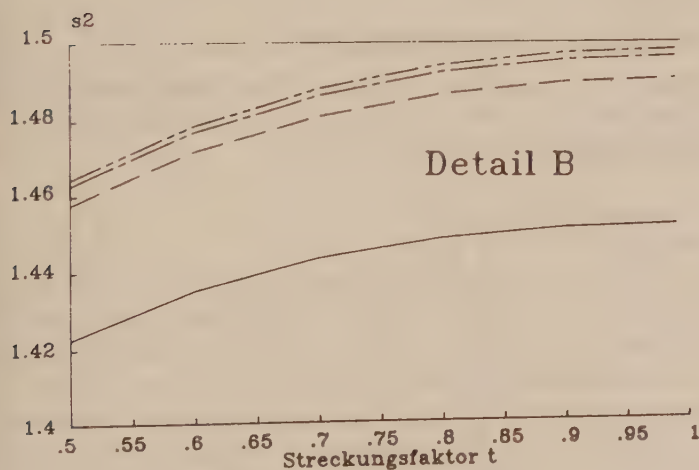
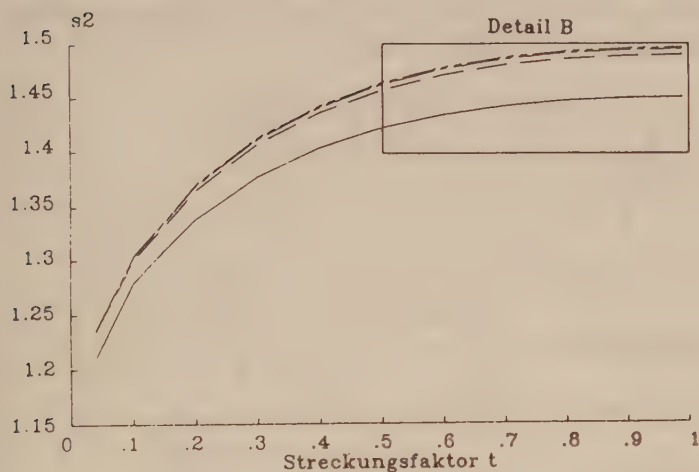
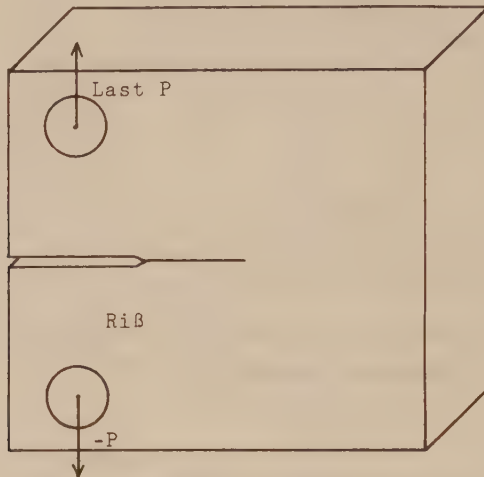


Bild 4: Konvergenz der aus den 2. Eigenwerten bestimmten Sing.-
ordnung gegen den wahren Wert von $s_2=1.5$.

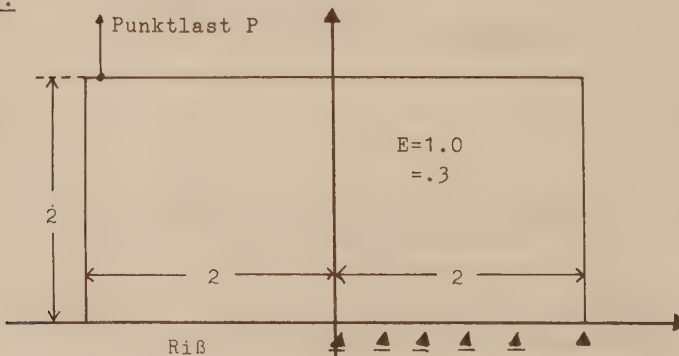
Als anwendungsorientiertes Beispiel aus der ebenen Elastizitätstheorie wurde die Spannungsverteilung an der Rißspitze einer Compact-Tension Probe untersucht. Es handelt sich dabei um eine genormte Standardprobe mit einem definierten Riß zur Bestimmung des sogenannten K_{IC} -Wertes, eines wichtigen bruchmechanischen Kennwertes (siehe z.B.: ROSSMANITH (1982)).

Skizze:



Aufgrund der Symmetrie genügt es dabei die obere Hälfte zu betrachten. Wir legen das folgende einfache Modell zugrunde mit E-Modul $E=1$. und POISSON-verh. $\nu = .3$.

Modell:



Bei der wahren Lösung hat der Vektor der Verschiebungen $\begin{pmatrix} u \\ v \end{pmatrix}$ eine Reihenentwicklung folgender Art (vgl. PAPAIOANNU, HILTON (1974)):

$$\begin{aligned} u(r, \omega) &= \sum_{i=1}^{\infty} k_i r^{i/2} f_i(\omega) \\ v(r, \omega) &= \sum_{i=1}^{\infty} k_i r^{i/2} g_i(\omega) \end{aligned} \quad (12.1)$$

mit:

$$\begin{aligned} f_1(\omega) &= F_1(\omega) \cos(\omega) - G_1(\omega) \sin(\omega) \\ g_1(\omega) &= F_1(\omega) \sin(\omega) + G_1(\omega) \cos(\omega) \end{aligned}$$

Die ersten 3 dieser Funktionen lauten dabei:

$$\begin{aligned} F_1(\omega) &= -(5/2-4h) \cdot \cos(\omega/2) + .5 \cdot \cos(3\omega/2) \\ F_2(\omega) &= (2-4h) + 2 \cos(2\omega) \\ F_3(\omega) &= (3/2-4h) \cdot \cos(\omega/2) + .5 \cdot \cos(5\omega/2) \\ G_1(\omega) &= (7/2-4h) \cdot \sin(\omega/2) - .5 \cdot \sin(3\omega/2) \\ G_2(\omega) &= -2 \cdot \sin(2\omega) \\ G_3(\omega) &= (9/2-4h) \cdot \sin(\omega/2) - .5 \cdot \sin(5\omega/2) \end{aligned} \quad (12.2)$$

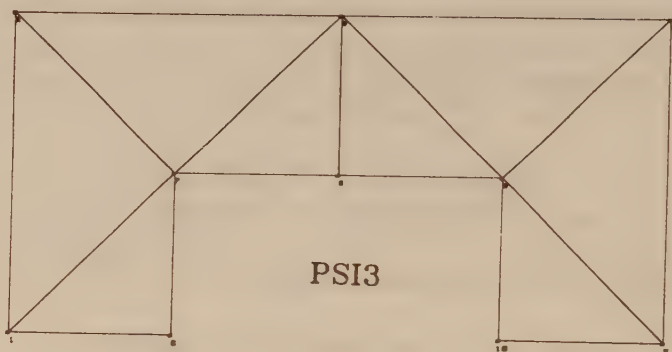
dabei ist h eine Materialkonstante, die sich aus dem POISSON-Verhältnis ν ergibt:

$$h = \begin{cases} \nu & \text{bei ebenen Verzerrungszustand} \\ \frac{\nu}{1+\nu} & \text{bei ebenen Spannungszustand} \end{cases}$$

Dieses Problem wurde mit der infiniten Substrukturtechnik mit verschiedenen Streckungsfaktoren $0 < t < 1$ und verschiedenen Elementarten gerechnet. Dazu wurde die Steifigkeitsmatrix der 1. Substruktur mit dem Finite-Elemente-Programm HP-FE-II der Firma HEWLETT-PACKARD auf einem (großen) Personal-Computer vom Typ HP9836C bereitgestellt. Diese Matrix wurde ausgekoppelt, entsprechend partitioniert und an die CYBER 845 (CDC) des Regionalen Rechenzentrums Erlangen (RRZE) übertragen. Dort wurden dann die latenten Wurzeln wieder über das aufgedoppelte Eigenwertproblem (9.3) mit Bibliotheksprogrammen der NAG-Library bestimmt.

Folgende 5 Zerlegungen und Elementarten wurden dabei verwendet:

Gitter 1



Gitter 2

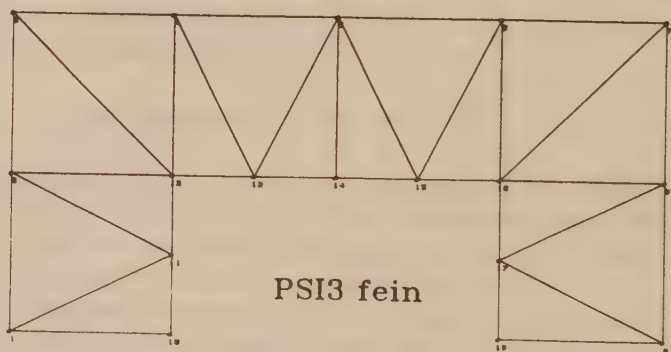
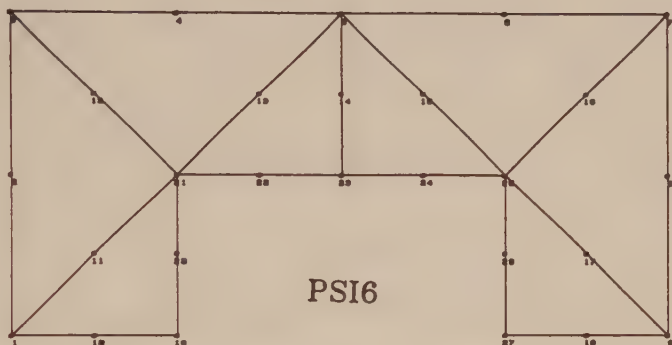


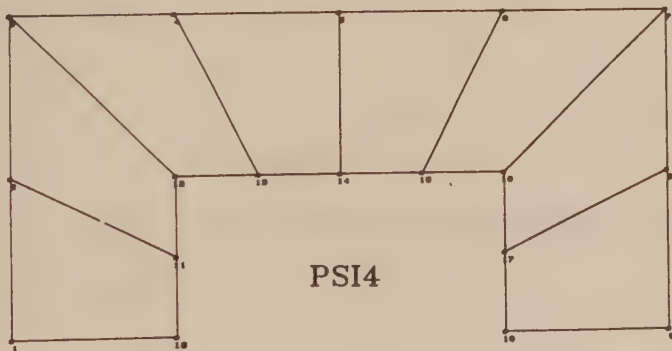
Bild 5: F.E.-Zerlegungen und äußerste Substruktur für die Anwendung bei der CT-Probe.

Gitter 3



PSI6

Gitter 4



PSI4

Bild 6: F.E.-Zerlegungen und äußerste Substruktur für die CT-Probe.

Gitter 5

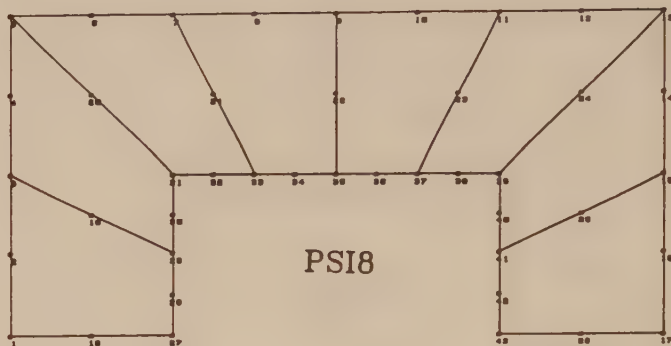


Bild 7: F.E.-Zerlegung der äußersten Substruktur bei der CT-Probe.

	Gitter 1	Gitter 2	Gitter 3	Gitter 4	Gitter 5
Elementtyp	PSI3	PSI3	PSI6	PSI4	PSI8
Knoten	10	18	27	18	43
Innenknoten	0	0	9	0	9
$\dim(A_\gamma), \dots$	10	18	36	18	52
lat. Wurz. = 0	1	1	19	1	19
lat. Wurz. = ∞	1	1	19	1	19
Streckungsf. t	.5	.5	.5	.5	.5

Tabelle 2: Übersicht über die verschiedenen Gitter.

Die folgenden Bilder zeigen jeweils die wahre Winkelverteilung, wie sie sich aus (12.1) auf den Knotenpunkten der Gitter ergibt und im Vergleich dazu die entsprechenden Eigenvektoren der Fixpunktmatrix X_∞ aus den Gittern PSI3 und PSI3-fein:

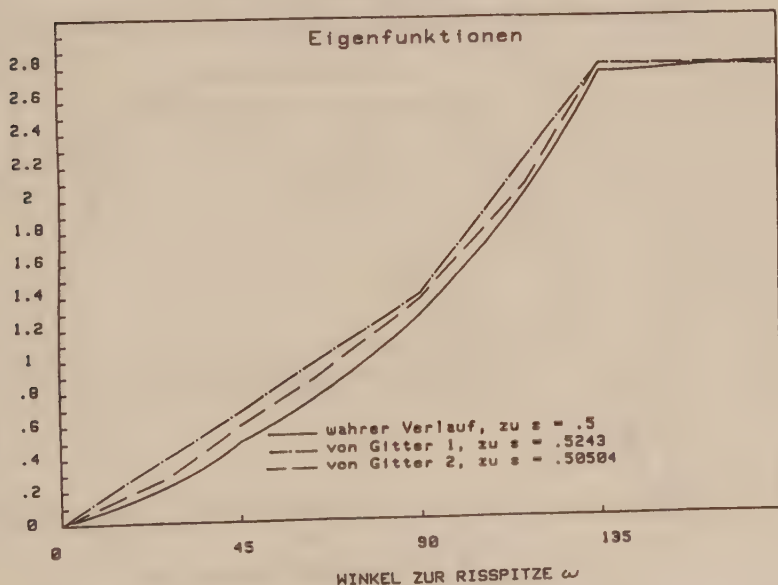
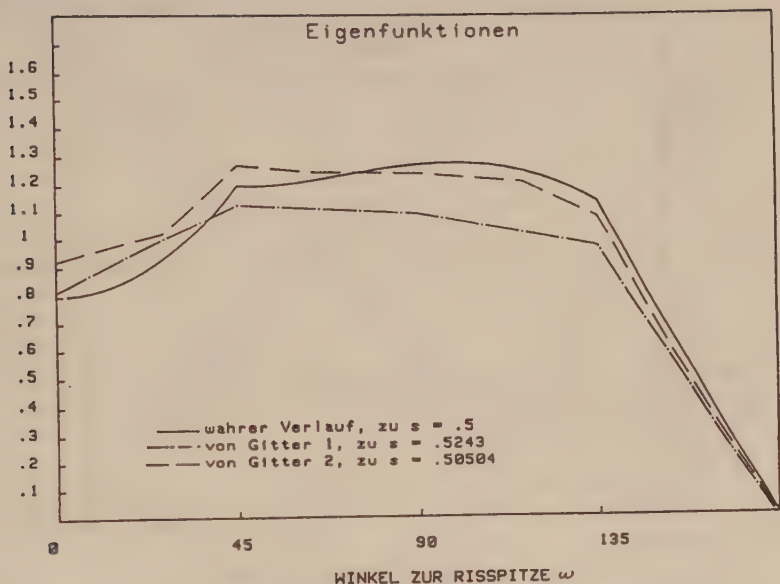


Bild 8: Vergleich der u bzw. v Verschiebungen des jeweils 1. Eigenvektors der Fixpunktmatrix X_{∞} mit den wahren Werten. (vgl. Formel (10.13) und Formel (12.1)) Oben: $f_1(\omega) \cdot r^{.5}$ und unten $g_1(\omega) \cdot r^{.5}$ auf den äußeren Gitterpunkten.

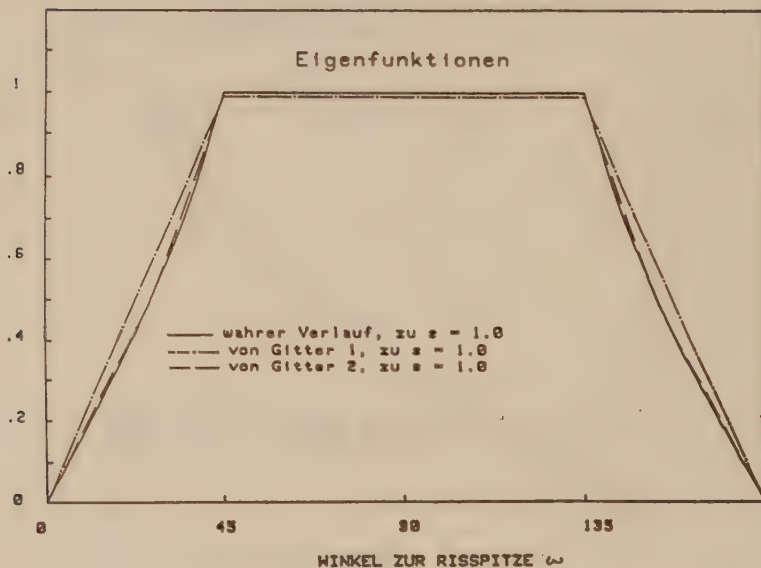
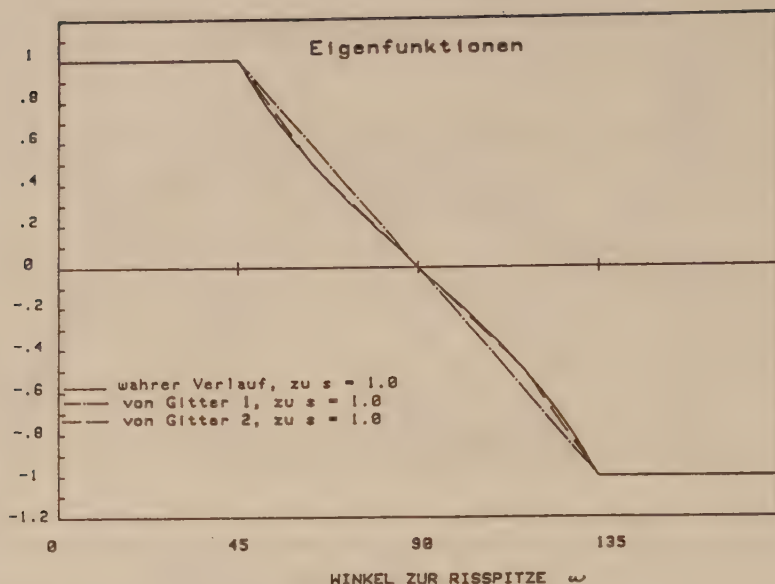


Bild 9: Vergleich der u bzw. v Verschiebungen des jeweils 2. Eigenvektors der Fixpunktmatrix X_∞ mit den wahren Werten. (vgl. Formel (10.13) und Formel (12.1)) Oben: $f_2(\omega) \cdot r^1$ und unten $g_2(\omega) \cdot r^1$ auf den äußeren Gitterpunkten.

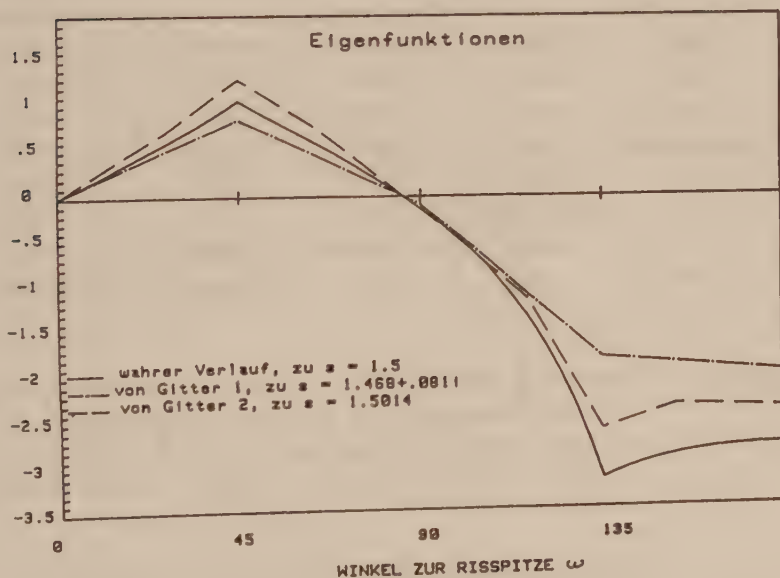
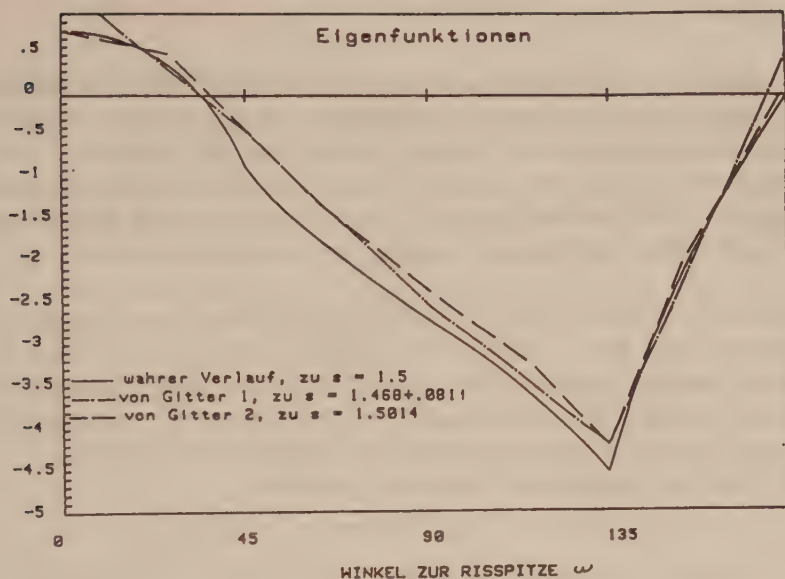


Bild 10: Vergleich der u bzw. v Verschiebungen des jeweils 3. Eigenvektors der Fixpunktmatrix X_∞ mit den wahren Werten. (vgl. Formel (10.13) und Formel (12.1)) Oben: $f_3(\omega) \cdot r^{1.5}$ und unten $g_3(\omega) \cdot r^{1.5}$ auf den äußeren Gitterpunkten.

In den Bildern 8,9,10 zeigt sich eine recht gute Übereinstimmung zwischen den Eigenvektoren der Fixpunktmatrix X_∞ und der Winkelverteilung der analytischen Lösung. Dabei ist zu beachten, daß Eigenvektoren nur bis auf eine Konstante eindeutig sind und somit die absolute Größe in den Bildern irrelevant ist. Die Skalierung wurde willkürlich nach graphischen Gesichtspunkten gewählt.

Besonders interessant ist auch das Auftreten komplexer Eigenwerte und Eigenvektoren bei allen Gittern. Diese scheinen stets dann zu entstehen, wenn die tatsächliche zugehörige Winkelverteilung nach dem Abtasttheorem nicht mehr auf den Gitterpunkten dargestellt werden kann. Die folgende Tabelle gibt jeweils die ersten 8 Eigenwerte zu den 5 gezeigten Gitterzerlegungen.

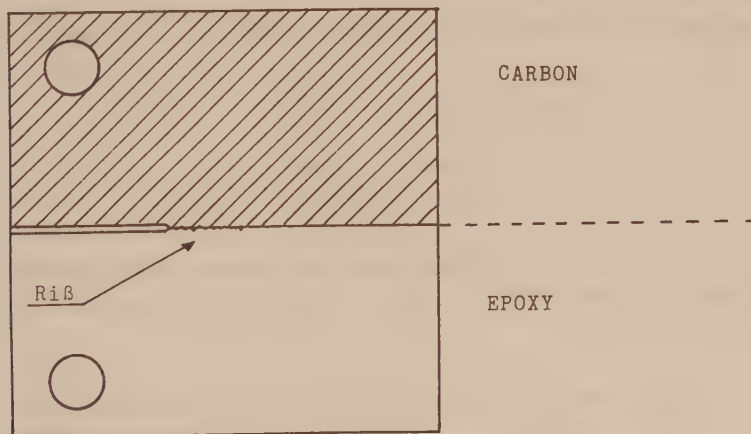
i	Gitter1	Gitter2	Gitter3	Gitter4	Gitter5
0	1.0000	1.0000	1.0000	1.0000	1.0000
1	.6953	.7046	.7073	.7120	.7073
2	.5000	.5000	.5000	.5000	.5000
3	.3608±.0201	.3532	.3534	.3457	.3534
4	.3042	.2794	.2500	.2274	.2500
5	.2153	.2057±.0241	.1760	.1625	.1765
6	.1029	.1702	.1014±.0391	.0930	.1299
7	.0537	.1682	.0917	.0372±.0591	.0837±.0091

Tabelle 3: Die ersten 8 Eigenwerte der Fixpunktmatrix X_∞ zu den 5 verschiedenen Gittern beim Streckungsfaktor $t=.5$. Die wahren Werte sind gegeben durch: $s_1=(.5)^{1/2}$ (1.0000, .7071, .5000, .3536, .2500, .1768, .1250, .0884)

Bei obiger CT-Probe sind die komplexen Eigenwerte wohl eher als numerischer Effekt, bedingt durch grobe Diskretisierung zu interpretieren. Prinzipiell kann ihr Auftreten aber nicht ausgeschlossen werden. Dazu folgendes Beispiel:

Bereits 1965 hat ENGLAND darauf hingewiesen, daß bei einem Riß, der auf die Grenzfläche zwischen 2 verschiedenen Materialien auftrifft, bei bestimmten Material- und Winkelkombinationen auch komplexe Singularitätenordnungen auftreten können. Die physikalische Interpretation solcher Terme der Form $r^s * f(\omega)$ mit komplexem s ist sehr problematisch. Bei LÖBEL, ZEILINGER, DESKA (1984) wird die Konsequenz für die Anwendung der bruchmechanischen Konzepte auf faserverstärkte Verbundwerkstoffe diskutiert.

Das vorgeschlagene Verfahren der infiniten Substrukturkondensation ist auch in diesen pathologischen Fällen anwendbar. Zur Demonstration betrachten wir hier folgende spezielle CT-Probe, deren eine Hälfte aus Kohlenstoff und die andere aus Epoxidharz ist. Der Riß möge dabei genau auf dem (perfect verklebten) Interface entlang laufen.



Durch Anwendung der MELLIN-Transformation hat BOGY (1971) gezeigt, daß sich die Singularitätenordnungen aus den Nullstellen der komplexen Gleichung

$$\sin^4(s\pi) * (\cos^2(s\pi) + b^2 * \sin^2(s\pi)) = 0 \quad (12.3)$$

ergeben. dabei ist b ein Bimaterialparameter, der sich aus den Schubmoduln G_1 , G_2 und den POISSON-Verhältnissen ν_1 , ν_2 der beiden Materialien berechnet:

$$b = \frac{G_1(m_2 - 2) - G_2(m_1 - 2)}{G_1 m_2 + G_2 m_1}$$

mit $m_i = \begin{cases} 4(1 + \nu_i) & \text{bei ebenem Verzerrungszust.} \\ 4/(1 + \nu_i) & \text{bei ebenem Spannungszustand.} \end{cases}$

Die Nullstellen von (12.3) findet man zu: $(i := \sqrt{-1})$

$$s_k = k/2 + 1/\pi \tanh^{-1}(b) \cdot i \quad k=1, 2, 3, \dots \quad (12.4)$$

und $s_j = j \quad j=1, 2, 3, \dots$

Legt man ebenen Spannungszustand zugrunde und die Werte:

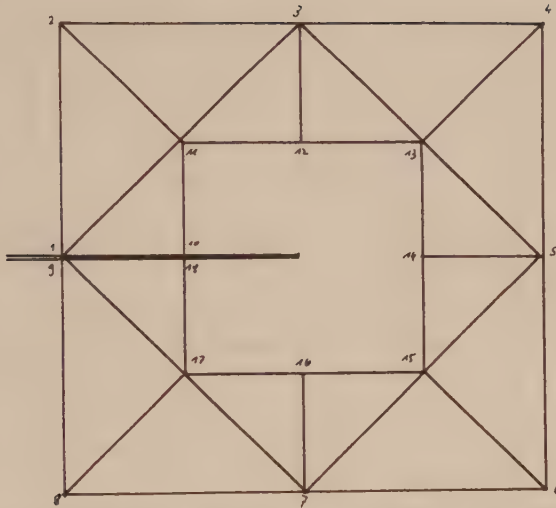
$$\begin{array}{ll} G_1 = 177.2727 & \} \text{ Carbon} \\ \nu_1 = .10 & \} \\ G_2 = .92596 & \} \text{ Epoxy} \\ \nu_2 = .35 & \} \end{array}$$

so ergibt sich $b = .3200637$ und die ersten 3 Singularitätenordnungen s haben damit die Werte (nach der Größe von $\text{Re}(s)$ sortiert):

$$\begin{aligned} s_1 &= .5 + .10559i \\ s_2 &= 1.0 \\ s_3 &= 1.5 + .10559i \end{aligned}$$

Dieses Problem wurde mit folgender 1. Substruktur und F.E.-Zerlegung mit der infiniten Substrukturkondensation gerechnet, wobei als Streckungsfaktoren $t=.5$ und $t=.95$ benutzt wurde:

(Skizze für $t=.5$)



18 Knoten
 36 Freiheitsgrade
 16 Elemente (PSI3)
 $\dim(X_\infty) = 18$

Dabei ergaben sich folgende Eigenwerte und daraus berechnete Singularitätenordnungen:

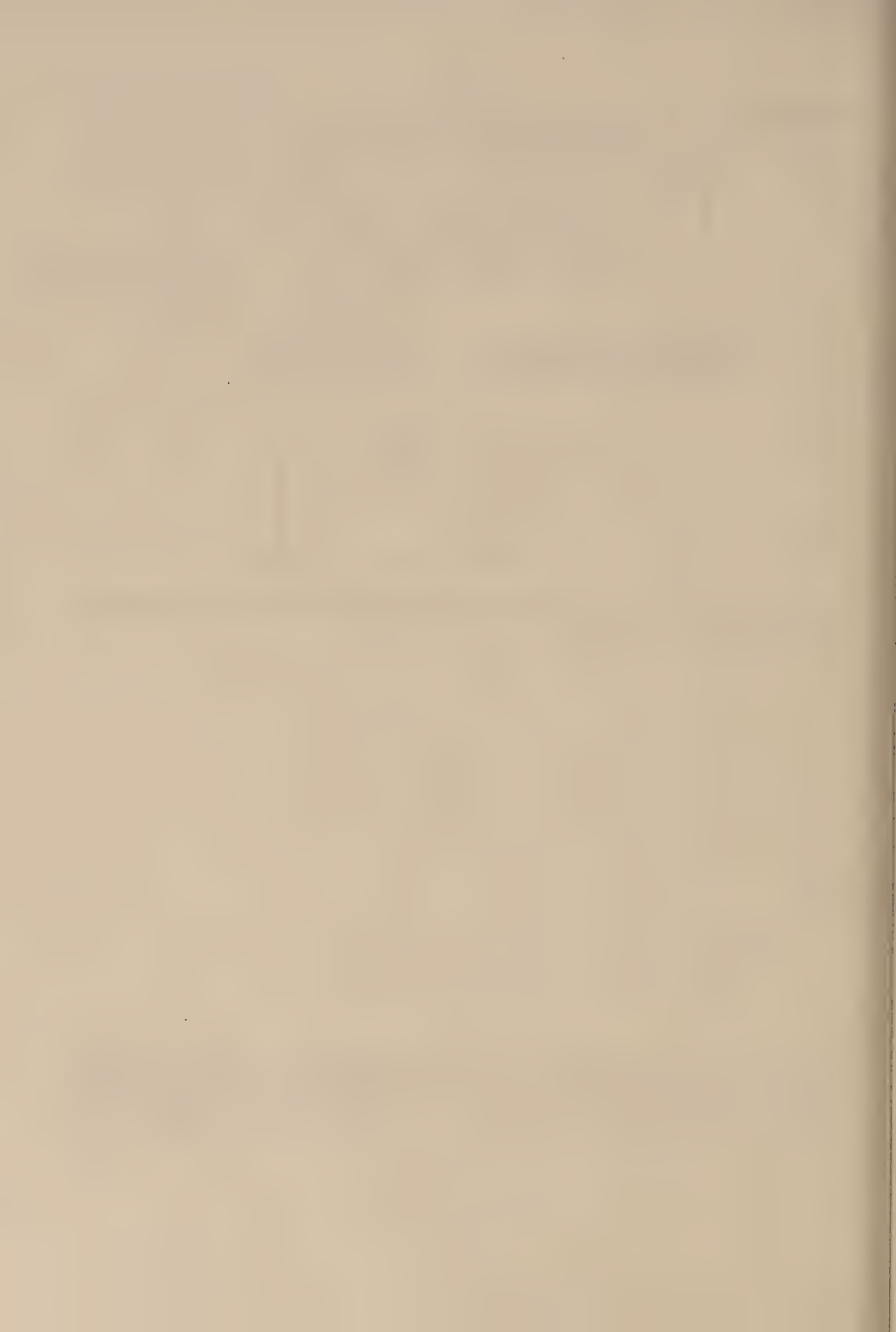
$t=.5$

Eigenwert	s		
$-.6924 + .03213i$	$.52878 + .06689i$	$\approx$	s_1
$-.5000$	1.00000	$\approx$	s_2
$-.3854 + .01871i$	$1.37377 + .06999i$	$\approx$	s_3

$t=.95$

Eigenwert	s		
$-.9735 + .00595i$	$.5226 + .11921i$	$\approx$	s_1
$-.9501$	$.99707$	$\approx$	s_2
$-.9339 + .00660i$	$1.33247 + .13784i$	$\approx$	s_3

Wie man sieht, ist bereits bei dieser sehr groben Diskretisierung die Zuordnung numerischer Wert - wahrer Wert ganz brauchbar. Jedenfalls wird mit diesem Beispiel plausibel, daß der Algorithmus prinzipiell die Möglichkeit komplexer Eigenwerte enthalten muß.



LITERATUR

- AFOLABI M.O.: An algorithm for inversion of a matrix with polynomial entries. Congressus Numeratum Vol. 34 (1982) pp 159-171
- ADAMS R.A.: Sobolev Spaces, New York u.a., Academic Press (1975)
- AZIZ A.K., DORR M.R., KELLOG R.B.: A new approximation method for the Helmholtz-equation in an exterior domain. SIAM J. Numer. Anal. 19,5. p 899-908, (1982)
- BABUSKA I.: The finite element method for infinite domains. J. Math. Comp. 26, p1-11, (1972).
- BARSOUM R.S.: On the use of isoparametric finite elements in linear elastic fracture mechanics. Int. J. num. Meth. Engng. 10 p25-37, (1976)
- BARSOUM R.S. : Triangular Quarter-point elements as elastic and perfectly plastic crack tip elements. Int. J. num. Meth. Engng 11, p85-98, (1977)
- BAYLISS A., GUNZBERGER M., TURKEL E.: Boundary conditions for the numerical solution of elliptic equations in exterior regions. SIAM J. Appl. Math. 42, p430-450, (1982)
- BEER : Infinite domain elements in finite element analysis of underground excavations. Int. J. Numerical and analytical Meth. in Geomechanics 7, p1-7 (1983)

ANHANG

- BEER, MEEK: Infinite domain elements. Int. J. num. Meth. Engng. 17, p43-52, (1981)
- BETTES P.: Infinite Elements. Int. Journ. Num. Meth. Engng. 11, p53-64, (1977)
- BOGY D.B.: Edge-bonded dissimilar orthogonal elastic wedges under normal and shear loading. Trans. ASME/E 35 pp460, (1968).
- BOGY D.B.: Two edge-bonded elastic wedges of different materials and wedge angles under surface tractions. Trans. ASME/E 38, pp377 (1971a).
- BOGY D.B.: On the plane elastostatic problem of a loaded crack terminating at a material interface. Journal of applied mechanics - Trans. ASME/E 38 p911-918, (1971b)
- BOGY D.B., WANG K.L.: Stress singularities at interface corners in bonded dissimilar isotropic elastic materials. Int. J. Solids & Structures (1971).
- BYSKOV E.: The calculation of stress intensity factors using the finite element method with cracked elements. Int. J. Fracture 26, p329-337, (1984)
- CANTOR M.: Numerical treatment of potential type equations on R^n Theoretical considerations. SIAM J. Numer. Anal. 20, p72-85, (1983)
- CIARLET Ph. G.: The finite element method for elliptic problems. Amsterdam (1978), North Holland Publ. Comp.
- CURNIER A.: A static infinite element. Int. J. Num. Meth. Engng. 19, p1479-1488, (1983)
- DAVIS G.J.: An algorithm to compute solvents of the matrix equation $AX^2 + BX + C = 0$. ACM Trans. Math. Software 9, Algo-

ANHANG

rithm 598 (1983)

DENNIS J.E., TRAUB J.F., WEBER R.P.: The algebraic theory of matrix polynomials I, SIAM J. Numer. Anal., Vol 13, p831-845, (1976),

DENNIS J.E., TRAUB J.F., WEBER R.P.: The algebraic theory of matrix polynomials II, SIAM J. Numer. Anal., Vol 15, p523-533, (1978)

ENGLAND A.H.: A crack between dissimilar media. Trans. ASME/E 32 p400-410, (1965)

FENNER D.N.: Stress singularities in composite materials with an arbitrarily oriented crack meeting an interface Int. Journ. of Fracture 12, p705-721 (1976)

ERDOGAN F., GUPTA G.: The stress analysis of multilayered composites with a flaw. Int. J. Solids & Structures 7, pp39 , (1971).

GALLAGHER R.H.: A review of finite element techniques in fracture mechanics. In: Numerical Methods in Fracture Mechanics edited by Luxmore, Owen. (1979).

GANTMACHER F.R.: The theory of matrices I, II, Chelsea, New York (1960).

GOLDSTEIN C.I.: The finite element method with nonuniform mesh sizes for unbounded domains. Math. of Computations 36, p387-404, (1981a)

GOLDSTEIN C.I.: The finite element method with nonuniform mesh sizes applied to the exterior Helmholtz problem. Numerische Mathematik 38, p61-82, (1981b)

ANHANG

- GOLDSTEIN C.I.: A finite element method for solving Helmholtz type equations in waveguides and other unbounded domains. Math. of Computation 39, p309-324, (1982)
- HENSHELL R.D., SHAW K.G.: Crack tip elements are unnecessary. Int. J. num. Meth. Engng. 9, p495-507, (1975)
- HEPPLER G., HANSEN J.S.: Mixed mode fracture analysis of rectilinear anisotropic plates by high order finite elements. Int. J. f. Num. Meth. Engng. 17, p445-464, (1981).
- HOUE HAN : The numerical solutions of interface problems by infinite element method. Numerische Mathematik 39, p39-50, (1982)
- IRWIN : Handbuch der Physik (ed. Flügge), Vol 6, p551, Springer Berlin (1958)
- JOHNSON A.R.: On the accuracy of polynomial finite elements for crack problems. Int. Journ. Num. Meth. Engng. 17, p1835-1842, (1981)
- KARP S.N., KARAL F.C.: The elastic-field behaviour in the neighborhood of a crack of arbitrary angle. Comm. pure appl. Math. 15, p413-422, (1962)
- KELLOG R.B.: Singularities in interface problems. in: Numerical solutions of PDE - II, SYNSPADE (1971), p351-400
- KELLOG R.B.: On the POISSON-equation with intersecting interfaces. Applicable Analysis 4, p101-129, (1975).
- KONDRATJEV V.A.: Boundary problems for elliptic equations in domains with conical or angular points. Trans. Moscow. Math. Soc., Vol 16, p227-313, (1967)

ANHANG

- LANCASTER P.: Lambda-matrices and vibrating systems. Pergamon Press Oxford London NewYork Paris, (1966)
- LEHMAN R.S.: Developments at an analytic corner of solutions of elliptic partial differential equations. Journal of Math. and Mech. Vol 8, p727 (1957)
- LEIS R.: Vorlesungen über partielle Differentialgleichungen zweiter Ordnung. BI-Taschenbuch Mannheim, (1967).
- LIONS, MAGENES: Problem aux limites non homogenes et application. Vol. 1, Dunod Paris (1968)
- LÖBEL G., ZEILINGER H., DESKA R.: Theoretische Grundlagen für die Anwendung der Bruchmechanik auf Faserverbundwerkstoffe. Zeitschrift für Werkstofftechnik 15, 277-287, (1984)
- LYNN, HADID: Infinite elements with $1/r$ type decay. Int. J. num. Meth. Engng. 17, p347-355 (1981)
- McDUFFEE C.: The theory of matrices. Chelsea New York, (1946).
- MOLER C.B., STEWART G.W.: An algorithm for generalized matrix eigenvalue problems. SIAM J. Numer. Anal. 10, p241-256, (1973)
- MOTZ H.: The treatment of singularities of part. diff. equations by relaxation methods. Q. Appl. Math. 4, p371-390 (1946)
- PAPAIOANNU S.G., HILTON P.D.: A finite element method for calculating stress intensity factors and its application to composites. Engng. Fracture Mech. 6, pp807-823, (1974).
- ROSSMANITH H.P.: Grundlagen der Bruchmechanik, Springer Wien, New York (1982).

ANHANG

- SARANEN J.: Zur Approximation der Strahlungslösungen mit der Methode der finiten Elemente. ZAMM 61, T256-T257, (1981)
- SCHATZ A.H., WAHLBIN L.B.: Maximum norm estimates in the finite element method on plane polygonal domains. I
Mathematics of computation , Vol 32 p73-109 (1978)
- SCHATZ A.H., WAHLBIN L.B.: Maximum norm estimates in the finite element method on plane polygonal domains. II .
Mathematics of computation , Vol 33 p465-492 (1979)
- SCHWARZ H.R.: Methode der finiten Elemente. Teubner Studienbücher Bd. 47, Stuttgart (1980)
- SHAO XIU MING, WANG SHU JING: The FOURIER-method for solving systems of equations in the infinite similar elements method. Numer. Math. J. Chinese Univ. 5, (1983), p160-169, (Chinese with english summary).
- STAAB G.H., SUN C.T.: Singular finite element efficiency in estimating stress intensity factors. Int. Journ. Num. Meth. Engng. 17, p557-572 (1981)
- STOER, BULIRSCH: Einführung in die numerische Mathematik I/II. Heidelberger TB 105/114, Berlin, Heidelberg, New-York (1973)
- STRANG G., FIX G.J.: An analysis of the finite element method. Englewood cliffs, N.J. (1973)
- TAYLOR C., HOOD P.: A numerical solution of the Navier-Stokes equations using the finite element technique. Computers and Fluids 1, p73-100 (1973)
- THATCHER R.W.: Singularities in the solution of Laplace's equation in two dimensions. J. Inst. Math. Applics. 16, p303-319, (1975)

ANHANG

- THATCHER R.W.: The use of infinite grid refinements at singularities in the solution of Laplace's Equation.
Numerische Mathematik 25, p163-178, (1976)
- THEOCARIS P.S.: The order of singularity at a multi-wedge corner of a composite plate. Int. J. Engng. Sci. 12, p107, (1974).
- WALSH : The computation of stress intensity factors by a special finite element technique. Int. J. Solids and Structures Vol 7, p1333-1342, (1971)
- WATWOOD V.B.: The finite element method for prediction of crack behaviour. Nuclear Engineering and Design 11, p323-332, (1969)
- WILLIAMS : Stress singularities resulting from various boundary conditions in angular corners of plates in extension.
Trans ASME/E 19, (1952), p526
- WILSON W.K.: Finite element methods for elastic bodies containing cracks. in: Mechanics of Fracture (ed. Sih) Noordhoff Netherlands (1973), p484-515
- WLOKA J.: Funktionalanalysis und Anwendungen. De Gruyter, Berlin (1971)
- ZIENCIEWICZ, EMSON, BETTES: A novel boundary finite element. Int. J. num. Meth. Engng. 11, p55-64, (1977)
- ZIENCIEWICZ O.C.: The finite element method. McGraw-Hill Book Company (UK) Limited, 3. Auflage London (1977)
- ZURMÜHL R.: Matrizen und ihre Technischen Anwendungen. Springer Verlag Berlin, Göttingen, Heidelberg (1964).

SYMBOLVERZEICHNIS

Symbol	Seite	Bedeutung
Ω	8	Grundgebiet
$a(.,.)$	8	Bilinearform
f	8	Funktional
V	8	Hilbertraum
V_h	9	endl. dim. Teilraum aus V
u_h	9	Näherungslösung aus V_h
ψ_i	9	Basis von V_h
TRIM3/PSI3	11	3-eckiges Element mit lin. Ansatz
TRIM6/PSI6	11	3-eckiges Element mit quad. Ansatz
PSI4/PSI8		4/8-eckiges Element mit lin/quad. Ansatz
L	22	Kondensationsoperator
a^*, f^*, u_f	22	kondensierte Ersatzaufgabe
$K U=F$	9	lin. Gleichungssystem der F.E.-Methode
$\mathcal{T}$	33	Isomorphismus in V
γ	33	Verträglichkeitsfaktor
$T, \hat{T}$	33	affine Abbildung
$R_0, R_1, \dots$	45	Substrukturen
N	45, 48	Anzahl der verschachtelten Substrukturen
A_Y, A_O, A_u	51, 53	bestimmte Teile der partitionierten Steifigkeitsmatrix u. rechten Seite von der O. Substruktur
B	51, 53	
C	51, 53	
F	51, 53	
U_i	53	Unbekannte Knotenwerte der i. Substrukt.
K_i	56	Rekursionsmatrizen
X_i	56	
H_i	56	
S_O	48, 49	Steifigkeitsmatrix der O. Substruktur
$\mathcal{T}$	60	Rekursionsoperator "v. unten nach oben"
$\mathcal{G}$	61	Rekursionsoperator "v. oben nach unten"

Symbol	Seite	Bedeutung
D	61	Definitionsbereich von $\mathcal{F}$
Γ	63	Einheitskugel in D
κ	64	Kontraktionsmodul von $\mathcal{F}$ (bzw. $\mathcal{G}$)
$\rho(\dots)$	62	Spektralradius einer Matrix
λ		Eigenwert, latente Wurzel
v		Eigenvektor, latenter Vektor
X_∞	70	Fixpunkt von $\mathcal{F}$
Y_∞^j	70	Fixpunkt von $\mathcal{G}$
X_i	88	Zwischenschritte zu X_∞
r, ω	90, 1	Polarkoordinaten
t	91	Streckungs-/Stauchungsfaktor
$k_i, \tilde{k}_i$	90	Entwicklungskoeffizienten
s_i	90	Singularitätenordnungen
f_i	90	Winkelfunktionen (allg.)
$W^{s,p}(\Omega)$	97	SOBOLEV-Raum
$H^1(\Omega)$	97	"
$\ \cdot \ _{s,p}$	97	SOBOLEV-Norm
$ \cdot _{s,p}$	97	Semi-Norm
$\mathcal{T}_{\Omega,h}$	96	Interpolationsoperator

LEBENS LAUF

Gustav Ernst LÖBEL geboren am 12.05.53 in Landsberg/Lech
Eltern: Gustav und Gerda Löbel
eine Schwester (geb. 1965)

Schulen Sept. 1959 Volksschule Kaufering
Sept. 1964 Oberrealschule Landsberg/Lech
08.06.1973 Abitur

Wehrdienst 02.07.73 - 30.09.74 als Sanitäter/Luftwaffe

Studium ab Wintersemester 74/75 Mathematik, Neben-
fach Physik an der Ludwig-Maximilians-Uni-
versität München.
16.05.80 Diplom

Tätigkeiten ab Juni 1980 freier Mitarbeiter bei der
Firma M.A.N. München Abt. Neue Technologie
(Materialprüfung, Faserverbundwerkstoffe)

seit 1.1.84 wissenschaftl. Mitarbeiter am
Lehrstuhl für Technische Mechanik der Uni-
versität Erlangen- Nürnberg (Boundary-Ele-
ment-Methoden)

Familienstand 23.09.77 Heirat mit Kristine Schallermeir
22.03.83 Sohn Benjamin

